اثر طول الاختبار على تحديد درجة القطع لاختبار تحصيلي محكي المرجع في مادة علم نفس الشخصية

أ. م.د. إيمان عبد الكريم ذيب

الجامعة العراقية/ كلية التربية

الفصل الاول

التعريف بالبحث

مشكلة البحث والحاجة اليه

وجه اصحاب الاتجاه المعاصر في القياس انتقادات حادة الى النظرية التقليدية في القياس النفسي والتربوي والاساليب القائمة بموجبها اهمها عدم استقلالية نتائج القياس عن العينة المطبق عليها الاختبار وعن الفقرات المتضمنة في الاختبار ، مما دعا الى ظهور الاتجاه المعاصر في القياس النفسي والتربوي بهدف التوصل الى اعلى مستوى من الدقة والموضوعية في القياس بحيث يحقق ادق علاقة بين اداة القياس والسمة الكامنة لدى الفرد (Nunnally ,1978,316).

تعتبر الاختبارات محكية المرجع أكثر انواع الاختبارات التحصيلية ملاءمة لقياس وتقويم تحصيل الطلاب ، وذلك لأنها تقوم على تحديد المهارات والكفايات المطلوب إتقانها بدقة فائقة لكي يتمكن المعلم من قياسها وملاحظتها بشكل مباشر ومن ثم تقدير مدى ما حققه الطالب من تلك الاهداف ، بناء على مستوى اداء محدد ، وهذا مما يساعد المعلم على عملية التشخيص ، حيث يتم تصنيف الطلاب الى فئتين متقنة وغير متقنة للمهارات والكفايات المحددة ، يلي ذلك تحديد نقاط الضعف والقوة في مستوى تحصيل الطالب ومعرفة المهارات التي اتقنها والتي لم يتقنها ، وبذلك يستطيع المعلم وضع البرامج العلاجية المناسبة لذلك ،وتعتبر خطوة تحديد درجة قطع الاختبار محور ارتكاز بناء الاختبارات محكية المرجع ، فمن خلالها تتحقق الافتراضات الاساسية التي يقوم عليها القياس محكي المرجع ،فهي تزود المعلم بتقديرات كمية لتحديد مدى تمكن الطالب

من المهارات والحكم على ادائه ومن ثم اتخاذ القرارات المناسبة (علام ١٤،١٩٨٦-١٥)،لذلك فان عملية تحديد درجة قطع الاختبار تؤثر تاثيراً مباشراً على تلك القرارات التربوية ، فأي خطأ ينتج عن تحديد هذه الدرجة يؤدي الى الحصول على نتائج سلبية (علام ٨٦،١٩٨٦-٩١)، وتعتبر عملية تحديد درجة القطع في الاختبارات محكية المرجع من قضايا القياس والتقويم الهامة التي اختلفت فيها رؤى علماء القياس والتقويم التربوي والنفسي ، حيث كثر الجدل والنقاش بينهم حول تصميم الطريقة المثلى لتحديد درجة القطع ، لذلك حاولوا تقديم الأسس التي بموجبها يتم تحديد درجة القطع المناسبة التي يمكن الوثوق بها في تصنيف الطلاب ، وقد حرصوا في هذه المحاولات على تلافي خطا التداخل بين الفئتين المتقنة وغير المتقنة للمهارات التي يقيسها الاختبار والتي عادة ما تنتج عن اخطاء القياس ومن تلك الاخطاء (Berk,1976,5) ،عامل الصدفة وعامل الغش وعامل التخمين ، وعادة ما تؤدي هذه العوامل الى ظهور خطأي التصنيف الاول(α) (تصنيف الطالب المتقن ضمن فئة غير المتقنين) والثاني (β) (تصنيف الطالب غير المتقن ضمن فئة المتقنين) والتي يترتب عليهما نتائج وأضرار سلبية على الطالب ، وهي اعادة الطالب المتقن الوحدة الدراسية مرة أخرى بالرغم من تمكنه من المهارات المطلوبة ، وفي المقابل تمكين الطالب غير المتقن بالوحدات الدراسية اللاحقة (عبد السلام ،١٩٩٢ ،١٢٠-١٢١)، وقد تعددت طرق تحديد درجة القطع باختلاف رؤى علماء القياس التربوي والنفسي ، الامر الذي أدى الى اختلاف قيم درجات القطع باختلاف رؤى علماء القياس التربوي والنفسي ، عندتطبيق تطبيق الطرق المختلفة على اختبار موحدوعلى عينة

دراسية موحدة ،وذلك ما أكدته دراسات هالين وآخرون ، وسكاكون وكلينج ، وآندرو وجميس (Haplin et al, 1983,185-196;Skakun&Kling,1980,229-235 ;Andrew&Hect,1976,45-50).

وبالرغم من ان من ان معظم الدراسات السابقة التي قارنت طرق تحديد درجات القطع المختلفة أكدت على وجود فروق جوهرية ، الإ ان ما هو متوفر في أدبيات البحث العربي لا يكفي لتوفير القاعدة العلمية التي يمكن أن يعتمد عليها مصممو الاختبارات في اختيار الطريقة التي تتلاءم مع طبيعة الاختبار وطبيعة القرارات التعليمية المرتبطة به .

وبناءً على ما تقدم تحاول الدراسة الحالية معرفة أثر طول الاختبار على تحديد درجة قطع الاختبار الناتجة عن تطبيق أربعة من طرق تحديد درجة القطع اثنتين تحكيميتان وهما : طريقتي أنجوف ونيدلسكي واثنتين تعتمدان جزيئاً على التحكيم وتسترشدان ببيانات تجريبية وهما طريقتي المجموعات المتضادة والمجموعات المحكّية.

في ضوء ما سبق تحدد مشكلة الدراسة في التساؤل الرئيس التالي :

هل تختلف درجة قطع الاختبار الناتجة عن طرق تحديد درجة القطع المختلفة نيدلسكاي –أنجوف –المجموعات المتضادة – المجموعات المحكية عند تطبيقها على اختبار موحد وعينة دراسية واحدة بأختلاف طول الاختبار ؟

وتأسيساً على ماتقدم يمكن أن تظهر أهمية الدراسة الحالية فيما يلي :

١) تتناول هذه الدراسة بالبحث الاختبارات التحصيلية محكية المرجع التي كثر فيها النقاش والجدل بين علماء القياس والتقويم التربوي والنفسي ،والتي تحتاج الى مزيد من العناية والاهتمام

والبحث ، وذلك نظراً لحداثة هذا النوع من القياس مقارنة مع الانواع الاخرى من المقاييس التربوية والنفسية .

٢) تأتي أهمية الدراسة من الحاجة الماسة لتبصير التدريسي بأهمية تحديد درجة القطع في الاختبارات محكية المرجع ، لأنها يعّول عليها الحكم على ما يستطيع وما لا يستطيع أن يتقنه الطالب ، ومن ثم اتخاذ القرارات التربوية المناسبة حول تمكين الطالب المتقن من الفصول الدراسية اللاحقة ، وإعادة تأهيل الطالب غير المتقن للمهارات المراد قياسها في الاختبار .

٣) تنبع أهمية الدراسة من أهمية تحديد درجة القطع في تصنيف الطلاب الى فئتين متقنة وغير متقنة للمهارات المطلوبة ، يليه تشخيص تلك المهارات وتحديد نقاط الضعف والقوة ، ومن ثم وضع الخطط العلاجية المناسبة .

٤) تسهم هذه الدراسة في توضيح طرق تحديد درجة القطع الاختبارات محكية المرجع ، بهدف مساعدة المعلم على اختيار الطريقة المناسبة لأداته ، وذلك لتعدد هذه الطرق واختلاف افتراضاتها واهدافها ونتائجها ، وحيث أن صدق القرارات التربوية يعتمد على الاختيار والتطبيق المناسبين لطريقة تحديد درجة القطع ، لذا فإنه يمكن القول بأن دراسة طرق تحديد درجة القطع أمر في غاية الأهمية وذلك لأنها تسهم في إيضاح العلاقة بين اختيار الطريقة والهدف من تطبيق الاختيار ،بالإضافة الى أن التعرَف على هذه الطرق سيعطي المعلم مؤشرات حول نقاط الضعف والقوة لهذه الطرق ، الامر الذي يساعد على ازالة الحيرة عند اختيار طريقة تحديد درجة القطع .

٥) قد تكون أول دراسة تطبّق على طلبة الجامعات العراقية –على حد علم الباحثة –والتي تهتم بالتعّرف على أثر طول الاختبار على طرق تحديد درجة القطع في الاختبارات محكية المرجع .

أهداف البحث :

يستهدف البحث الحالي تعرف اثر طول الاختبار على تحديد درجة القطع لاختبار تحصيلي محكي المرجع وتطلب تحقيق هذه الاهداف الاتي:

١- بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الشخصية . - ٢-تعرف اثر طول الاختبار على تحديد درجة القطع لاختبار علم نفس الشخصية وفق اربعة طرق كلا على حدة :

أ-طريقة أنجوف

ب-طريقة نيدلسكي

جـ-طريقة المجموعات المتضادة

د-طريقة المجموعات المحكية.

٣-تعرف اثر طول الاختبار على تحديد درجة القطع لاختبار علم نفس الشخصية وفق الطرق المستخدمة في الدراسة مجتمعة .

حدود البحث :

يقتصر البحث الحالي على:

-الحدود العلمية : تعرف اثر طول الاختبار على تحديد درجة القطع

-الحدود البشرية : طلبة كلية التربية –الجامعة المستنصرية من كلا الجنسين الذكور والاناث .

-الحدود الزمانية :العام الدراسي ٢٠١٠-٢٠١١

-الحدود المكانية :كلية التربية / الجامعة المستنصرية

تحديد المصطلحات :سيتم تحديد المصطلحات الآتية :

١-**طول الاختبار** Test Length : وهو عدد الفقرات المكونة للاختبارات الفرعية التي قامت الباحثة بتقسيمها في هذا البحث الى ثلاثة انواع كالتالي:

أ- اختبار طويل يتكون من (٤٢) سؤال

ب-اختبار متوسط يتكون من (٢٨) سؤال

جـ-اختبار قصير يتكون من (١٤) سؤال .

٢- درجة القطع Cut off Score :

عرّف هامبلتون درجة القطع بأنها " نقطة على متصل درجات الاختبار تستخدم لتصنيف الطلاب الى فئتين تعكس مستويات الأداء المختلفة بالنسبة لهدف معين (أو الاهداف) المراد قياسها في الاختبار "

كما عرّف بيرك Berk درجة القطع المثلى بأنها " تلك الدرجة التي تزيد من نسبة التصنيفات الصحيحة او في المقابل تقلل من نسبة التصنيفات الخاطئة " (Berk,1982,99) .

كما عرّفها بافام Popham " بأنها مقياس لمدى ملائمة أداء الطالب لهدف محدد " (Halpin et al, 1983, p.185) .

اما التعريف الإجرائي لدرجة القطع في هذه الدراسة فهو عبارة عن الدرجات الناتجة من تطبيق الطرق المختلفة في تحديدها والتي تقابل نقطة المحك المحددة مسبقاً وبدقة على متصل التحصيل والتي تصنف المتعلمين الى متقنين وغير متقنين للمحكوك الدراسي المحدد بدقة والذي يقيسه الاختبار .

اما الطرق المستخدمة في الدراسة فهي:

أ- طريقة نيدلسكاي Nedlesky's – Method :

أشار نيدلسكاي (1954) Nedlesky الى ان هذه الطريقة تستخدم مع الاختبار من نوع الاختبار من متعدد ، وتتطلب من كل محكم فحص كل مفردة من مفردات الاختبار ثم يحدد من البدائل الخاطئة لكل مفردة تلك التي من الممكن أن يتجنب اختيارها الطالب ذو الحد الأدنى من الكفاية ، ويكون الحد الأدنى لاحتمال الإجابة الصحيحة عن المفردة هو مقلوب عدد البدائل الباقية (Meskanska,1976,135).

ب-طريقة أنجوف Angoff's –Method (1976)

ذكر Angoff (1976) بأن هذه الطريقة تستخدم أيضاً مع الاختبار من نوع الاختيار متعدد وتتطلب هذه الطريقة من كل محكم التنبؤ بمستوى صعوبة كل مفردة في الاختيار وذلك عن طريق فحص كل محكم لمفردات الاختبار ، ثم يطلب منهم تصور مجموعة من الطلاب من ذوي الحد الأدنى من الكفاية ،ثم يقدر احتمال ان يعرف الطالب ذو الحد الأدنى من الكفاية المناسبة إجابة المفردة بصورة صحيحة دون أن يلجأ الى التخمين ، ثم يضع دائرة حول الرقم الذي يمثل مقدار الاحتمال المبين في ميزان فترى ، ثم تجمع القيم الاحتمالية المقدرة لكل مفردة ، ويمثل متوسط مجموع هذه الاحتمالات درجة القطع(Angoff,1971,9).

جـ-طريقة المجموعات المتضادة Contrasting Group

أشار هامبلتون Hambleton (1982) بأنه يتعّين على المحكمين في هذه الطريقة تحديد مجموعتين من الطلاب الأولى متقنة والثانية غير متقنة للمهارات التي يقيسها الاختبار ، وبعد تطبيق الاختبار يتم رسم منحنيين يمثلان التوزيع التكراري لدرجات الطلاب ،

وتمثل نقطة تقاطع المنحنيين درجة قطع الاختبار ، ويمكن تعديل الدرجة للتقليل من خطأ التصنيف الأول (α) وخطأ التصنيف الثاني(β) ،(Berk,1984,178-179).

د- طريقة المجموعات المحكية Criterion Groups

ذكر هامبلتون (1982) Hambleton بأن هدف هذه الطريقة هو تحجيم أخطاء التصنيفات الخاطئة وزيادة التصنيفات الصحيحة ، وتشابه هذه الطريقة الى حد ما طريقة المجموعات المتضادة الإ أنها تتطلب إيجاد التصنيفات الثنائية للاحتمالات التالية (طلاب متقنين فعلاً –طلاب غير متقنين فعلاً –طلاب متقنين غير حقيقيين – طلاب غير متقنين غير حقيقيين) يلي ذلك تحديد معامل صدق درجات القطع المحتملة ، ثم حساب المنفعة والضرر النسبيين لأخطاء التصنيف(Berk,1984,183) .

٣- الاختبار المحكي المرجع Criterion referenced Test :

يعرفه بافام على انه الاختبار الذي يستخدم لتقدير اداء الفرد في نطاق سلوكي محدد تحديدا دقيقاً (Popham1978,931) .

ويعرفه جلاسرونيتكو بأنه :

ذلك الاختبار الذي يعد يهدف احداث قياسات يمكن تفسيرها بشكل مباشر بناء على معايير (مستويات) اداء محددة (Gray ,1978,p.226) .

اما التعريف الاجرائي فهو اختبار تحصيلي في مادة علم النفس الشخصية وهو يتكون من (١٢٧) فقرة موزعة على (٩) اختبارات فرعية يقيس كل منها كفاية معينة من الكفايات الخاصة بمادة علم نفس الشخصية .

الفصل الثاني

الإطار النظري

تمهيد:

يقصد بالاختبارات محكية المرجع (CRTS) والتي يقصد بها أن مستوى المفحوص عند نقطة ما على متصل التحصيل ، أو هو أنواع السلوك التي يقوم بها خلال أدائه الاختبار ، أو هي الدرجة التحصيلية لاداء مرغوب فيه عند أي مستوى محدد ، وبذلك أصبحت لها تسميات عدة منها اختبارات التفوق (proficiency Test) ، أو اختبارات الاتقان (Competency Test)، أو اختبارات المهارات الأساسية (Best Skills Test) (الشرقاوي وزملاؤه، ١٩٩٦: ٨٩).

كما يذكر " بافام " (Popham) بأن الاختبار المحكي المرجع هو الاختبار الذي يستخدم للتاكد من مستوى الفردبالنسبة الى مجال سلوكي محدد بصورة جيدة ، للتاكد من مركز الفرد بالنسبة الى مجال سلوكي محدد(Popham، 1978 ، p.94) . وبهذا فان الاختبارات محكية المرجع تختلف عن معيارية المرجع بان نقطة الارتكاز تحدد غالبا عند اعلى مستوى للاداء الصحيح لبعض القدرات والمهارات المعنية . فالدرجة التي تكون في اعلى الاختبار تدل على مستوى االتمكن الكامل ، او على مستوى للاداء الصحيح لبعض المهارات او القدرات المعنية ، والدرجة التي تكون في اسفل الاختبار ، فانها تدل على ادنى مستوى لهذة القدرة ، وفيما يلي شيء من التفصيل عن تلك الاختبارات .

انواع الاختبارات محكية المرجع

أ- الاختبارات مجالية المرجع : Domain-referenced Tests

هي نوع من الاختبارات محكية المرجع التي تتطلب تحديدا دقيقا ومفصلا للمعارف والمهارات المكونة للنطاق السلوكي الذي يقيسة الاختبار ، حيث يذكر علام (٢٠٠١) بانها الاختبارات التي تتطلب تحديد مواصفات تفصيلية للنطاق السلوكي الشامل ، يمكن انتقاء عينة عشوائية ، او عشوائية طبقية من المكفردات التي يتم سحبها من النطاق الشامل ، والتي تقيس محتوى هذا النطاق المحدد بدقة (علام ٢٠٠١أ :٩٥) . اما نادية عبد السلام تعرفة بانة الاختبار المكون من عينة عشوائية او عشوائية طبقية من المفردات التي تمثل عينة شاملة من المفردات والتي تقيس مهارات معينة (عبد السلام ، ١٩٩٦ : ٩٤) . وبهاذا فان التحديد الدقيق والمفصل للنطاق السلوكي المتمثل بعينة عشوائية او طبقية عشوائية من المفردات ، فانها تساعد في تعميم نتائجالاختبار على كامل النطاق .

د-الاختبارات هدفية المرجع : Objective-referenced Tests

يذكر الشرقاوي وزملاؤه بان الاختبار هدفي المرجع والاختبار محكي المرجعهو شيئا واحد بشرط عدم الاكتفاء بان الاختبار هدفي المرجع هو الاختبارالذي يقوم على اساس مجموعة من الاهداف المحددة والواضحة وانما لابد من التحديد الدقيق والمفصل لمستوى تلك الاهداف (الشرقاوي وزملاؤه ، ١٩٩٦ :٢٩) في حين بين علام (٢٠٠١)بانها الاختبارات التي تبنى على اساس مجموعة الاهداف التعليمية المصاغة صياغة اجرائية (سلوكية) (علام ، ٢٠٠١ أ :٩٦) . أما " بافام " (Popham) فاستخدمها في قياس مجموعة من الاهداف التعلمية الاجرائية وذللك لتمتعها بقلة عدد مفرداتها والتي تمثل النطاق السلوكي المراد قياسه (Popham 1978 ,P 93) وبهذا فان هذة الاختبارات لها ميزات منها اقتصادية وتحدد مستوى الاداء لكل هدف ، وواقعية ، فضلا عن انها تتيح للمعلمين تطبيقها

في الوقت المخصص للحصة الدراسية ، هذا ماجعل الباحث يختار هذا النوع من الاختبارات بغية تحقيق اهداف بحثة .

استخدامات الاختبارات محكية المرجع :

١- لمعرفة المتقدمين الذين اكتسبو القدرات والمهارات التي من المتوقع اكتسبها بعدعملية التعليم .

٢- تحديد موضع المتعلم بالنسبة لمستويات اداء محددة .

٣- لتقدير أو تقويم المهارات الاساسية للمتعلم لتحديد أدائه في مجال تعليمي ما .

٤- تقويم برامج التدريس والمناهج .

٥- لمعرفة فيما اذا كان المتعلم قد حصل على الحد الادنى أملافي المحتوى المقاس .

٦- تحديد الاهداف التي حصل عليها كل متعلم على حدة .

٧- تشخيص نواحي القصور في التحصيل .

٨- للتاكد من التعلم لاشياء مخددة . (علام ، ٢٠٠١ أ : ٢٢) (الشرقاوي وزملاؤه ، ١٩٩٦ : ٦٩-٧٠) (أبو علام ، ١٩٨٧: ١٣٣) ابراهيم ، ١٩٩٠:٢٤)

بناء الاختبارات التحصيلية محكية المرجع

يمر بناء الاختبار محكي المرجع في جانب التحصيل بمراحل عدة هي :

اولا-تحديد المحتوى المراد قياسة :

اذا كان المحتوى المراد قياسة محدود فانة يمكن الاكتفاء بمعرفة مكونات هذه الوحدة ، اما اذا كان المحتوى عريضا ومتسعا فيمكن تقسيمه الى موضوعات فرعة مرتبطة بعضها البعض بحيث يمكن قياسها كوحدة واحدة (علام ، ١٩٨٦ : ٣٧) ، ويتم التحديد عن طريق مايلي :

ا-تحديد الكفايات الاساسية المراد تحقيقها من خلال تدريس هذا المحتوى : ويتم ذلك عن طريق :

١- الاستعانة بمجموعة صغيرة من خبراء المادة الدراسية .

٢- تحليل المحتوى المراد قياسة الى نواتجة السلوكية .

٣- اجراء دراسة لتقدير حاجة الفئة المستهدفة .

بشرط ان تكون الكفاية أعلاه محدودة الاتساع وقابلة للاكتساب والتنمية ، وان ينتقل اثر تعلمها في مجال محتواها العام الى مجالات الاخرى ، وان تمثل السلوك الختامي في البنية الهرمية (علام ، ٢٠٠١ أ : ٤٢) .

ب-تحليل الكفايات الرئيسية الى مكوناتها الاساسية :

في هذه الخطوة يتم تحليل الكفايات المركبة تحليلا اجرائيا او بنائيا ثم ترتيب مكوناتها ترتيبا منطقيا باحدى الطرق الاتية :

١- **طريقة تحليل الاجرائات Procedural Approach :** تميز هذة الطريقة في تحليل الكفايات بان تعلم كل خطوة يعتبر مدخلا للخطوة التالية بمثابة سلسلة خطية متتابعة من الخطوات المستقلة والتي تسهم في النهاية الى تحقيق الكفاية الرئيسية المراد قياسها .

٢- **طريقة التحليل الهرمي Hierarchical Approach:** تستخدم هذه الطريقة عندما يكون من غير الممكن تعلم الكفاية الرئيسية دون تعلم مهارات فرعية تسبقها وهنا يتم تحديد المعارف والمهارات التي ينبغي تعلمها بترتيب هرمي .

٣- **طريقة تجمع بين تحليل الاجرائات والتحليل الهرمي :** Combination Approach:تستخدم هذة الطريقة عندما يكون هناك كفايات معقدة تحتاج الى تحليل اجرائي ذا خطوات مستقلة متتابعة تتضمن احكاما متعددة ، وتحليلا هرميا يتضمن شبكة من المعارف والمهارات المساعدة والتي تؤدي في النهاية الى تحقيق تلك الكفايات المعقدة (علام ٢٠٠١ أ : ٤٢).

ج- صياغة الاهداف السلوكية :

في الخطوة السابقة تم تحديد الكفايات المراد قياسها والتي تمثل المحتوى المراد قياسة ايضا ، ومن ثم تحليل تلك الكفايات بواسطة عدة طرق تم ذكرها اعلاه الى مكوناتها باكثر دقة ، اما في هذة الخطوة يتم صياغة نواتج عملية التحليل صياغة اجرائية يمكن ملاحظتها بشكل مباشر وقياسها بمفردات اختيارية وهي بمثابة اهداف سلوكية تمثل الكفايات الرئيسة التي تم تحليلها (علام ، ١٩٨٩ : ٣٩) وكالاتي:.

١- **بناء مفردات الاختبار :**

يتم بناء مفردات الاختبار محكي المرجع في الجانب الحصيلي بمرحلتين هما :

أ- **تحديد مواصفات الاختبار** :ان الاهداف الاجرائية التي تم تحديدها بالخطوة السابقة لاتكون واضحة بدرجة تسمح بتحديد نطاق مفردات الاختبار التي تقيس السلوك المطلوب ، اذ يقترح " بابام " (Popham) طريقة مواصفات الاختبار التي تشتمل على اربعة مكونات اساسية بالاضافة الى ملحق المواصفات ، وهي المكونات هي :

١- **الوصف العاما تقيسة المفردة :** هو الصياغة الاجرائية للهدف الذي تقيسة المفردة .

٢- **عينة من المفردات :** هو بناء عدد من المفردات التي تقيس الهدف المحدد وهي بمثابة فكرة واضحة ومحددة عن كيفية قياس الهدف .

٣- **عناصر المثير :** عبارة عن مجموعة من العبارات ، نحاول عن طريقها وصف وتحديد نوع المثيرات التي تقدمها مفردات الاختبار التي تقدمت مفردات الاختبار .

٤- **عناصر الاستجابة :** هي عبارة عن مجموعة من العبارات تحاول عن طريقها تحديد نوع وطريقة الاستجابة على مفردات الاختبار ، فاذاكانت

المفردة من نوع الاختبار من متعدد ، فان عناصر الاستجابة تتضمن قواعد تكوين المشتتات وطبيعة الاجابة الصحيحة .

٥- **ملحق المواصفات :** هنا تسجل بعض التفصيلات التي تؤدي الى مزيد مزيد من التوضيح لمحتوى مفردات الاختبار وعناصر المثيرات والاستجابات (Popham ,1978 ,P .121-128).

ب-كتابة مفردات الاختبار :

هنا يتم كتابة مفردات الاختبار التي يجب ان ترتبط ارتباطا وثيقا بالمحتوى المراد قياسة ، ويشترط في كتابة المفردات ان تميز بين التلاميذ الذين اقنوا الهدف والذين لم يتقنوه ، وقسم علام ١٩٨٦ هذه الخطوة الى ثلاث مراحل هي :

١- **اختيار انواع المفردات المناسبة لقياس الاهداف :** فمن المعلوم ان بعض انواع المفردات تصلح لقياس اهداف معينة افضل من غيرها ، فمثلا تصلح مفردات الاختبار من تعدد لقياس التذكر والفهم والتطبيق احيانا ، بينما تصلح مفردات المقال لقياس القدرة على تنضيم المعلومات والاستنتاج والتفسير وهكذا .

العلاقة بين عدد المفردات وبين الاهمية النسبة للنطاق السلوكي المراد قياسة ، اي يجب ان يناسب العدد الذي تشمتل علية عينات المفردات مع الاهمية النسبية للنطاق السلوكي الذي نقيسة .

أ- العلاقة بين عدد المفرداتوالحدى الادنى للثبات المطلوب .

ج- العلاقة بين عدد المفردات وزمن تطبيق الاختبار.

٢- **كتابة مفردات الاختبار :** ان مفردات الاختبار تبنى على اساس مواصفات النكاق السلوكي ، وان يكون مستوى صعوبة كل مفردة مناسبا لصعوبة الهدف الذي تقيسة ، وتكون عينة المفردات ممثلة للنطاق السلوكي ، بالاضافة الى مراعاة الاصول الفنية واللغوية في صياغة المفردات (علام ،١٩٨٦ ٤٠) .

الخصائص السيكومترية للاختبار التحصيلي محكي المرجع :

ان الاختبارات التحصيلية محكية المرجع بشكل عام :

١- تهتم بتقدير الفرد في ضوء محك اداء محدد .

٢- توزيق الدرجات في الاختبارات محكية المرجع تاخذ بشكل ملتوي التواء سالب ، خاصة اذا كان البرنامج التعليمي فاعلا .

اذن فهي لاتهتم بتحديد المكانة بالنسبة للفرد بين اقرانة وبالتالي لاتستبعد ايا من المفردات طالما انها تقيس هدفا سلوكيا مهما ، لذا فلاداعي لايجاد معامل صعوبة مفردات اداة الدراسة الحالية ولاحتى غيرها ، اضافة الى اعتماد انموذج راش في التحليل الاحصائي لفقرات الاختبار . وبالتالي لاضرورة للتكلم عنها في هذا الموضوع ، وانما يكتفي الباحث بالكلام عن صدق الاختبار وثباتة .

أ- صدق الاختبار محكي المرجع :

يشير الصدق الى مدى استخدام درجات المقياس في القيام بتفسيرات معينة كأن تفسير درجات اختبار تحصيلي ما ، على انها تمثل مجال التحصيل الذي يقيسة هذا الاختبار ، وهذا المفهوم يمكن تحقيقة عن طريق جمع الادلة باستخدام عدة طرق (علام ١٩٩١ : ٨٧-٨٨)وهناك ثلاث طرق مختلفة لصدق الاختبار المحكي المرجع وهي :

١- **الصدق الوصفي : Descriptive Validity**

يعد الاختبار المحكي صادقا اذا استطعنا باستخدامه وصف اداء الفرد بالنسبة للنطاق السلوكي الذي يقيسة ذلك يالاختبار ، واحيانا يشار له بصدق المحتوى حيث يساعد الباحث بان يصف اداء الفرد بالنسبة للنطاق السلوكي ، ويقدر هذا الصدق عن طريق احكام الخبراء في تطبيق المفردة للهدف وذلك باستخدام المعادلة الاتية :

(N-Lik ----------)

حيث ان :

lik=مؤشر تطابق المفردة بالهدف الخاص بالمفردة (K)للهدف (i).

N=عدد الاهداف

Xijk=عدد المحكمين المتفقين (i)على المفردة (k) تقيس الهدف (i) ولا تقيس غيرة من خلال التقديرات (

٢-الصدق الوظيفي : Functional Validity

ويقصد بهذا النوع من الصدق ان يؤدي الاختبار المرجع الى المحك الوضيفية التي صمم من اجل تحقيقها او الغرض الذي بني من اجلة ، ولا يقتصر فقط على وصف النطاق السلوكي الذي يقيسه ، وهذا النوع من الصدق يناظر مايعرف بالصدق التجريبي Empirical Validity في الاختبارات المرجعة الى المعيار ، والذي يدل على قدرة الاختبار على التنبؤ باداء الفردى في مواقف تختلف الى حد ما عن المواقف يقيسها الاختبار . ونظرا لان الاختبارات المرجعة الى المحك لها وضائف واستخدامات متنوعة تهم المربين وخبراء التقويم بعضها لايتضمن التنبؤ بمحكات معينة .

٣- صدق انتقاء النطاق السلوكي (التكوين الفرضي) : Domaain Selection Validity

وتتخلص فكرة هذا النوع من الصدق في كيفية صياغة قواعد يمكن الاسترشاد بها في تكوين المواصفات التي تحدد النطاق السلوكي لمفردات الاختبار . وهذا يؤدي الى السؤال التالي :

مامدى هدف النطاق السلوكي الذي يمكن اختيارة من بين النطاقات السلوكية الاخرى في تمثيل الابعاد التي تهتم بقياسها ؟ وبالتالي يكون هدف الباحث هو اختيار نطاق سلوكي من بين النطاقات التي نهتم بقياسها ؟ وبالتالي يكون هدف الباحث هو اختيار نطاق سلوكي من بين النطاقات السلوكية

الممكنة بحيث يسمح له بالتعميم على النطاق الشامل للبعد الذي يود قياسة (Popham,1978,143-145) .

ب- ثبات الاختبار المحكي المرجع :

الثبات هو مدى اتساق التصنيف بناء على الدرجات المستمدة منها ، وسوف يعرض الباحث اهم طرق تقدير ثابت الاختبار محكي المرجع ، حيث تتطلب احد (هذه الطرق تطبيق الاختبار نفسه على نفس العينة ، بينما يتطلب البعض الاخر تطبيق صورتين متوازيتين للاختبار ، على ان تشمل كل من الصورتين على عينة عشوائية من المفردات المستمدة من نطاق شامل للمفردات التي تقيس هدفا تعليميا معينا) .

١- الطرق التي تطلب تطبيق الاختبار على نفس العينة من الطلاب وهي :

أ- **طريقة ليفينجستون (١٩٧٢) : Livingston Index**

وقد اعتمد ليفينجستون في اشتقاقه لمعامل الثبات الخاص بالاختبارات المرجعة الى المحك (CRTs) على مفهوم درجة القطع والتي تفصل بين المتقنين وغير المتقنين .

وبهذا فان هذا المعامل يهتم (بايجاد تباين درجات الافراد عن مربع الفرق متوسط تلك الدرجات عن درجة القطع) (Sax & Newton ,1997 ,P.283)

ب- **طريقة الاتفاق لصابكوفياك (١٩٧٦) : Subkoviak Method of Agreement**

وتعتمد هذه الطريقة على تقدير مدى اتساق تصنيف الطلاب الى متقنين وغير متقنين باستخدام نتائج تطبيق الاختبار مرة واحدة فقط ، وتتخلص هذه الطريقة في تقدير احتمال تصنيف كل فرد تصنيفا صحيحا Pc(i) وجمع هذه الاحتمالات وايجاد متوسطها ، وبذلك نحصل على معامل التفاف خاص بالمجموعة التي يطبق عليها الاختبار .

٢- الطرق التي تتطلب تطبيق صورتين متوازيتين للآختبار وهي :

أ-طريقة كابا لسواميثان ،وهاملتون ،والجانيا (١٩٧٤)

Swaminathan ,Hambleton,and Algina Method

وهذا المعامل يعبر عن درجة اتساق تصنيفات الطلاب في مرتـــي تطبيق الاختبار ويعبر عنه بالصورة الرياضية التالية :

معامل كابا $$\frac{(k)=p-pc}{1-pc}$$

حيث ان :

Pc نسبة الاتفاق المتوقع في التصنيفات

P نسبة الاتفاق الملاحظ في التصنيفات

Pc=P1.P.1+P0.P.0

حيث ان :

.P1 : احتمال تصنيف الاتقان في أحد صورتي الاختبار

P.1: احتمال تصنيف الاتقان في الصورة الاخرى

.P0 احتمال تصنيف عدم الاتقان احد صورتي الاختبار

P.0 احتمال تصنيف عدم الاتقان في الصورة الاخرى

((Crocker&Algina,1986,200-201)

ب-معامل كارفر (١٩٧٠) Carver Method

يعد معامل كارفر من الطرق الاولـــى المقترحـــة لتقـــدير ثبـــات الاختبارات محكية المرجع والتي تتعلق بأتساق قرار التصنيف ، وتقوم هذه الطريقة على تطبيق اختبارين متوازيين على مجموعة واحد من الافـــراد ، يليه مقارنة نسبة عدد الطلاب الذين تم تصنيفهم ضمن المتقنـــين فـــي كـــلا الاختبارين فاذا كانت النسبتين متماثلتين او قريبتين من بعضها فبذلك يعتبـــر الاختبارين ثابتتين ويمكن ايجاد قيمة معامل كارفر كما في الجدول

الاختبار (ب)		
	غير متقن	متقن
الاختبار (أ) متقن	ب	أ
الاختبار (أ) غير متقن	ج	د

وبحسب معامل كارفر بأستخدام الصيغة التالية : أ+ج

ن

حيث ان :

أ=مجموع الطلاب المتقنين بالفعل

ج=مجموع الطلاب غير المتقنين بالفعل

ن= أ+ب+جـ+د

تحديد طول الاختبار محكي المرجع

ذكر كل من هامبلتون وايقنور((Hambleton &Eignor,1980 بأن هناك عاملين ينبغي وضعهما في الاعتبار عند تحديد اسئلة الاختبار وهما:

١- العلاقة بين عدد أسئلة الاختبار والاهمية الخاصة بكفاية معينة : تتفاوت أهمية الكفايات المراد قياسها في الاختبار ، لذا ينبغي ان يختار مصمم الاختيار عند قياسه كفايات متعددة أسئلة تتناسب مع أهمية النطاق المراد قياسه .

٢-العلاقة بين عدد اسئلة الاختبار ومستوى الحد الادنى المقبول لدرجة ثبات الاختبار : يستدعي ذلك تحديد قيمة الدقة المطلوبة

(Degree of Precision) في تقدير درجة الطالب في النطاق ، ويتم ذلك وفق الصيغية التالية :

$$\text{طول الاختبار} = \frac{٠,٢٥}{(\text{درجة الدقة المطلوبة})^{٢}}$$

حيث ٠,٢٥ مقدار ثابت .

واشار كل من هامبلتون وآخرون (Hambleton et al (1978)) بأن الهدف من تطبيق الاختبار محكي المرجع هو تصنيف الطلاب الى فئتين متقنة وغير متقنة ، لذا فإن تحديد طول الاختبار المناسب الذي يقلل من أخطاء التصنيف (Number of Classification errors) يتطلب تطبيق اختبار طويل يقلل من أخطاء القياس .

وأضاف كل من هالادين ورويد Haladyn & Roid بأن قيمة أخطاء التصنيف تقل بشكل عام عند زيادة طول الاختبار لكن النقص في قيمة الخطأ يكون ملحوظاً عند الانتقال من اختبار طوله (١٠) فقرات الى اختبار طوله (٢٠) فقرة ، واما النقص في قيمة الخطأ عند الانتقال من اختبار طوله (٢٠) فقرة الى اختبار طوله (٣٠-٤٠) فقرة يكون صغير نسبياً لدرجة قد لا تبرر استخدام فقرات تزيد عن (٢٠) فقرة (الأحمد ،١٩٩٢ م) .

مسميات وتعريف درجة القطع Definitions of Cut off score

مسميات درجة القطع

تتعدد مسميات درجة قطع الاختبارات بتعدد آراء وخبرات علماء القياس والتقويم التربوي في مجال الاختبارات محكية المرجع ، ومن تلك المسميات ما ذكرها هالبن ، وآخرون Halpin ,et al (1983) وهي كما يلي :

١-درجة القطع Cut off score.

٢-الحد الادنى من الكفاية Minimum Competency.

٣- درجة المحك Criterion Score.

٤-درجة الاختبار Passing Score.

٥- مستوى الحد الادنى من الكفاية Level of Minimal Competency.

طرق تحديد درجة القطع

تبنى خبراء القياس والتقويم التربوي العديد من الطرق والمحاولات لتحديد درجة القطع ، واختلفت تلك الطرق والاساليب بأختلاف خبرات ودرجات تأهيل وتخصصات هؤلاء الخبراء ، فهنالك عدة طرق متعددة لتحديدها والتي ذكر بيرك **Berk (1986)** انها بلغت ثمان وثلاثين طريقة مختلفة من حيث أسلوب بناءها وتطبيقها وتفسير نتائجها ، ويمكن تصنيفها الى ثلاثة تصنيفات كالتالي :

اولاً : الطرق التحكيمية Judgmental –method : تستند الطرق التحكيمية على آراء المتخصصين في مجال القياس او في محتوى الاختبار المراد قياسه ، وذلك لتحديد درجة القطع المناسبة ، حيث يسترشد بآرائهم في تحديد مستوى الطالب ذو الحد الادنى من الكفاية (تحديد عدد الأسئلة التي يجب أن يجيب عليها الطالب بصورة صحيحة) واضعين في اعتبارهم المتغيرات التي تؤثر في تحديد هذه الكفاية

مميزات الطرق التحكيمية

لكل طريقة من الطرق التحكيمية مميزات وعيوب خاصة بها وهنالك عيوب ومميزات تكاد تكون مشتركة بينها وهي كالتالي:

كما استرسل بيرك **Berk (1986)** بأن من مميزات الطرق التحكيمية ما يلي :

١- انها سهلة الاستخدام مقارنة بالطرق الأخرى .

٢- انها سهلة التطبيق والفهم .

٣-انها سهلة التفسير .

٤-لا تحتاج الى جهد ووقت كبيرين .

٥-انها سهلة الحساب ولا تعتمد على اساليب احصائية معقدة .

عيوب الطرق التحكيمية

اشار كل من بيرك ،وهامبلتون ، وايغنور (Hambleton &Eignor,1980:Berk ,1986) الى عيوب استخدام الطرق التحكيمية وهي :

١-لانها اختيارية (Arbitrary) وسميت كذلك لأن هذه الطرق تعتمد كلياً على آراء المحكمين وتلعب كفاية هؤلاء المحكمون دوراً بارزاً في الوصول الى المعيار المناسب لتصنيف الطلاب وتحديد بنود الاختبار التي يمكن لأقل الطلاب كفاءة الإجابة عليها بشكل صحيح ويعتقد بأن كون هذه الطرق اختبارية لا يعد سبباً كافياً لرفض هذه الطرق لأن هذه العملية الاختبارية ليست مطلقة ولكنها تعتمد على مجموعة من الأسس في تحديد مستويات الاداء .

٢-صعوبة تحديد الطلاب الأقل كفاءة .

٣-لا تعتمد على بيانات تجريبية فعلية لكي يسترشد بها المحكمون في تقدير احكامهم .

٤- الحصول على درجات قطع مختلفة عند تطبيقها على اختبار موحد وعينة دراسية واحدة

٥-ذاتية المحكمين في اختيار الحد الأدنى من الكفاية وذلك بسبب اختلاف تصوراتهم حول عملية الاتقان .

٦-تفاوت عملية الاتساق الداخلي في عملية التحكيم.

٧-صعوبة التنبؤ بأداء الطلاب في الاختبار .

على الرغم من وجود بعض العيوب في استخدام الطرق التحكيمية الا انه توجد بعض المعايير المطروحة في ادبيات القياس التربوي التي تحد من تأثير هذه العيوب حيث ذكر بيرك (١٩٨٦) Berk بعض المعايير نذكر منها مايلي :

أ-استخدام الطرق لتوحيد بيانات الاداء الفعلي .

ب-يجب ان تكون عملية تحكيم مفردات الاختبار متزامنة مع تطبيق الاختبار .

ج-الاهتمام بتدريب المحكمين.

د-تنظيم عملية تحكيم بنود الاختبار .

هـ-ايجاد الحلول المناسبة لتسوية الاختلافات في احكام المحكمين.

و-تصميم استمارات تنظم عملية التحكيم .

ز-تزويد المحكمين ببيانات مسبقة حول اداء الاختبار .

وتناول كثير من الباحثين الطرق التحكيمية وأسهبوا في ذكر تفصيلاتها وتطبيقاتها ومميزاتها وعيوبها ومن ضمن هؤلاء الباحثين علام ، وبيرك ، وجايقر (Berk,1986;Jaeger,1989; ;علام،١٩٩٥) ومن هذه الطرق ما يلي :

١-طريقة أيبل Ebel's Method

اقترح ايبل Ebelطريقة لتحديد درجة القطع تعتمد على تعليقات المحكمين على مفردات الاختبار وذلك من خلال اعتبار ان مواصفات المفردات لها بعدين هما بعد الصلة Relevance وبعد الصعوبة difficulty ويشتمل بعد الصعوبة على ثلاثة مستويات (سهلة ، متوسطة ، صعبة) ويشتمل بعد الصلة على اربعة مستويات (اساسية ، مهمة ، مقبولة ، غير مهمة) وبذلك يستخدم

المحكم اعداداً مختلفة من تصنيفات الصعوبة والصلة ، ثم يطلب من كل محكم ما يلي :

أ-تحديد الخلية المناسبة لكل مفردة من المفردات بناءً على توافق محتواها مع تصنيفات البعدين .

ب-تحديد النسبة المئوية لعدد الطلاب ممن لديهم الحد الأدنى من الكفاية الذين يجب أن يجيبوا عن كل مفردة بصورة صحيحة وذلك لكل خلية من الخلايا ، ثم يضرب عدد مفردات كل خلية في النسبة المئوية ، ثم يقسم مجموع نواتج جميع الخلايا على عدد المفردات ، وبالتالي نحصل على درجة القطع (Jeager,1989) .

المميزات

يحدد تصنيف معامل صعوبة البند بالاستعانة ببيانات الاداء ، وتضرب في كلا البعدين مما يسهل عمل المحكم .

العيوب

تناول بيرك (١٩٨٦) Berk عيوب هذه الطريقة حيث ذكر بأن أيبل لم يلتزم بالمواصفات التي تستخدم في كل من البعدين والتمييز بينهما ، لكي يسترشد بها المحكمين في عملية التحكيم ، لذلك ربما يعتمد كل محكم على مواصفات يحددها لنفسه مما يؤثر بلا شك في دقة التصنيف واتساقه ، كما أن هذه الطريقة اعتمدت على تحكيم المفردة ككل ولم تأخذ بعين الاعتبار بدائل الاجابات التي تشتمل عليها مفردة الاختبار من متعدد .

بالاضافة الى ان طريقة ايبل تحتاج تحكيم خاص يتمثل بنسب مئوية للمفردات وذلك لكل خلية عبر بعدي الصلة والصعوبة المستمرين ، فاذا لم تكن هنالك معايير خارجية يبنى عليها هذا الحكم فإنه يبدو انه سيكون حكماً اعتباطياً .

للحد من هذه العيوب قام كل من سكاكون وكلينق Skakun &Kling بإجراء تعديلات على طريقة ايبل بحيث اصبحت تعتمد هذه الطريقة على بعدين هما معامل الصعوبة ويشتمل على (سهل –متوسط الصعوبة –صعب جدا) ومعامل التصنيف ويشتمل على (واقعي –شامل –حل المشكلات) وتستخدم بيانات الاداء في البنود لتقدير تصنيفات الصعوبة ويحدد كلا من باني الاسئلة ولجنة الاختبار الاسئلة طبقاً لمستويات التصنيف ، بناءً على ذلك تضرب الاسئلة في الخلايا المكونة من المصفوفة ٣×٣ ثم يقوم المحكمين بالإشارة الى نسبة المفردات في كل خلية والتي يجب ان يجيب عليها الطالب المتقن اجابة صحيحة (١٩٨٦، Berk) .

وتكمن عيوب هذه الطريقة في عدم وضع عاملي الصلة والاهمية في الاعتبار ومن المحتمل ايضا ان يصعب على بعد التصنيف تحديد مناطق المحتوى المحدد (١٩٨٦، Berk) . اذ انها لا تضع في اعتبارها مستويات صعوبة البنود الواقعية ، ومن المحتمل أن يصعب في بعد التصنيف تحديد مناطق المحتوى المحدد (١٩٨٦، Berk).

للحد من هذه العيوب قام كل من سكاكون وكلينق Skakun &Kling بإجراء تعديلات على طريقة ايبل بحيث اصبحت هذه الطريقة تعتمد على بعدين ايضا وهما معاملي الصلة والتصنيف واصبح معامل الصلة ابعاد ثلاثة (مهمة جداً –مهمة – مقبولة) وهذه الابعاد مقترنة ببعد معامل التصنيف .

٢-طريقة انجوف Angoff's Method (1971) يطلب من كل محكم في هذه الطريقة فحص كل فقرة من فقرات الاختبار ، ثم يطلب منهم تصور مجموعة من الافراد الذين حققوا الحد الادنى

للكفاية التي يقيسها الاختبار ، ثم يقدر نسبة عدد الافراد الذين يحتمل ان يجيبوا اجابة صحيحة عن كل مفردة من مفرداته ، ويمثل متوسط هذه النسب الحد الادنى لمستوى الاجتياز في الاختبار (Angoff,1971).

المميزات

أ-سهلة التنفيذ

ب-سهلة الفهم

ج-انها تعتمد على تصور عدد من الافراد الذين حققوا الحد الادنى للكفاية المطلوبة بدلاً من تصور فرد واحد .

العيوب

يصعب على كثير من المحكمين تعريف الطلاب من ذوي الحد الأدنى للكفاية ، ويصعب ايضاً تقدير القيم الاحتمالية لبعض المفردات المصاغة في عبارات منفية ((Negatively)) او المتضمنة عمليات حسابية ((Computational)) (١٩٨٦، Berk).

للحد من هذه العيوب أشار كل من جايقر وكروس وآخرون(Jeager,1989 ;Cross et al.,1984) بأن انجوف (Angoff) أجرى تعديلات على طريقته وذلك لتسهيل عملية تقديرات قيم المحكمين الاحتمالية لأسئلة الاختبار حيث يطلب من كل محكم تقدير احتمال ان يجيب الطلاب ذوي الحد الادنى للكفاية كل سؤال دون ان يلجوا الى تخمين (Guessing) الاجابة ، وتسجل تقديراتهم الاحتمالية في استمارة خاصة تحتوي على مقياس عشري الفترات "Equal Interval scal" (١=صفر -٠,١٠) (٢=٠,١١-٠,٢٠) (٣=٠,٢١-٠,٣٠) (٤=٠,٣١-٠,٤٠) (٥=٠,٤١-٠,٥٠) (٦=٠,٥١-٠,٦٠) (٧=٠,٦١-٠,٧٠) (٨=٠,٧١-٠,٨٠)

(٩=٨١،-٠,٩٠) (١٠=٩١،-١,٠) ثم يطلب من كل محكم اختبار القيمة الاحتمالية التي تمثل تقديره الخاص بكل سؤال من اسئلة الاختبار ، ثم تجمع القيمة الاحتمالية التي قدرها كل محكم لكل مفردة ، ثم نوجد متوسطها لتمثل درجة القطع .

٣-طريقة نيدلسكاي Nedlesky's Method

اقترح نيدلسكي (١٩٥٤) Nedlesky هذه الطريقة لتحديد درجة القطع للاختبارات التي تشتمل على اسئلة ذات اختبار من متعدد حيث تقوم مجموعة من المحكمين المختصين بتحديد درجة القطع المناسبة للاختبار ، وفي هذه الطريقة يطلب من كل محكم فحص كل مفردة من مفردات الاختبار ، ثم يحدد من بين الخيارات الخاصة لكل مفردة تلك التي يمكن ان يتجنب الطلاب ذوي الحد الادنى للكفاية اختبارها لأنها لا تمثل في نظرهم الاجابة الصحيحة للمفردة ، بعد ذلك يضع المحكم درجة للسؤال بناءً على ما تبقى من الخيارات ، وذلك بأخذ مقلوب الخيارات المتبقية .

فمثلاً اذا كانت المفردة تشتمل على خمسة خيارات ورأى المحكم ان الطالب ذو الحد الادنى للكفاية يمكن ان يتجنب اختيار ثلاثة منها عندها تكون درجة السؤال عبارة عن :

$$\frac{١}{٢} = \frac{١}{٥-٣} = ٠,٥$$

بعد ذلك يقوم المحكم بجمع درجات الاسئلة ، ويكون المجموع الكلي هو درجة القطع بالنسبة لذلك المحكم ، وتجمع في النهاية الدرجات التي يتحصل عليها المحكمون ، ثم يحسب المتوسط وبالتالي نحصل على درجة القطع للاختبار ككل ، غير ان ندلسكاي وجد انه من الافضل ايجاد الانحراف المعياري (σ) لدرجات القطع ثم ضربها في مقدار ثابت (K) يتفق عليه

المحكمون وتعتمد قيمة هذا المقدار على النسبة المئوية التي يرى المحكمون أنها مطلوبة لنجاح الطلاب من ذوي الحد الادنى للكفاية ، ويتم تحديد درجة القطع في هذه الطريقة وفق المعادلة التالية :

$$\overline{MPL} = \overline{MFD} + K\sigma FD + K\sigma FD$$

$$\overline{MFD} = \frac{\sum \overline{FD}}{N}$$

$$\sigma FD = \sqrt{\frac{\sum (MFD - \overline{MFD})^2}{N}}$$

حيث ان : MPL تمثل درجة القطع .

MFD تمثل متوسط الحد الادنى المحدد من المحكمين .

K تمثل المقدار الثابت المحدد من المحكمين .

σFD تمثل الانحراف المعياري لدرجات القطع .

N تمثل عدد المحكمين .

وأشار بيرك (Berk ،١٩٨٦) الى بعض عيوب هذه الطريقة وهي كالتالي :

لايوجد تبرير منطقي يؤيد افتراض (ندلسكاي ان الطالب الذي لا يعرف الاجابة الصحيحة لسؤال معين يلجأ الى التخمين العشوائي لأن مشتتات الأسئلة تصمم عادة لجذب غير المتمكنين ، كما ان هذه الطريقة تسمح فقط بقيم احتمالية منفصلة تعتمد على عدد خيارات الاجابات ،لذلك يضطر المحكمون لتعيين القيمة (٠,٥) لمعظم المفردات ، هذا بالاضافة الى ان هذه الطريقة تحتاج لقدر كبير من التدريب لكي يستخدمها المحكمون استخداماً مناسباً.

٤-طريق أنجوف –ندلسكاي المختلطة

Angoff Nedelsky combination ,Reid , 1984

في هذه الطريقة يتم حساب المعيار على اساس حساب متوسط المعايير الناتجة عن استخدام طريقتي انجوف ونيدلسكاي .

المميزات

لها نفس مميزات طريقة انجوف.

العيوب

لها نفس عيوب طريقتي انجوف ونيدلسكاي بالاضافة الى احتمالية زيادة مصادر الخطأ المختلفة وذلك عندما نحسب متوسط المعيارين (Berk ،١٩٨٦).

٥-طريقة تقدير أهمية الصعوبة difficulty importance estimate cangelosi ,1984

في هذه الطريقة يقوم المحكمون بفحص كل هدف ، وذلك لتقدير الحد الادنى للنجاح ويجب ان تتفق عينة المحكمين على الحد الادنى وعلى كل هدف ايضا ، وذلك لان للحد الادنى وزن خاص طبقاً لاهمية صلته بالهدف المنشود ، ثم تضرب النسبة المئوية لدرجة النجاح في نسبة اهمية الدرجة لكل هدف ، والناتج يجمع من خلال جميع الاهداف ، ويضرب في عدد بنود الاختبار وذلك للوصول للمعيار ، ويجب ان يضع الحكام في الاعتبار الخطأ المعياري وذلك عند اتخاذهم القرارات الخاصة بدرجة القطع.

المميزات

أ–سهلة الفهم .

ب–سهلة الاستخدام وتعمل على ربط درجة القطع بقائمة خصائص النطاق السلوكي .

العيوب

استخدمت تقديرات المحكمين للاهداف لتقدير معيار بنود الاختبار ، بينما نجد انه من المفروض ان تكون العلاقة بين الاسئلة والاهداف مباشرة ، وليس بالضرورة افتراض ان تقديرات الاسئلة تساوي تقديرات الاهداف ، ومن المحتمل ان يكون نوع الخطأ المعياري المستخدم غير ملائم لتفسير درجة القطع (١٩٨٦، Berk).

٦-طريقة مواصفات المفردة (ميلز وبار١٩٨٣)

في هذه الطريقة يقوم المحكمون بقراءة مواصفات كل مفردة والتي تحتوي على المهارة المراد قياسها ووصف محتوى البند المقبول وخواص المثير والقوانين التي تحكم احداث خيارات الاستجابة وعينة المفردات ، ومن ثم فإن على المحكمين ادراك وتصور الاطار الكلي للمفردات والمتمثل في المواصفات والتنبؤ بنسبة البنود في الاطار الكلي ، والتي يمكن ان يجيب عليها الممتحن ذو الحد الادنى من الكفاية اجابة صحيحة .

المميزات سهلة الفهم وسهلة الاستخدام وتقوم بربط درجة القطع بمواصفات النطاق السلوكي وربطهما بكافة البنود التي تقيس النطاق السلوكي وتستغرق زمناً اقل مما لو عمدنا الى تقدير كل مفردة بشكل منفصل ، وتكون درجة القطع فيها مرتبطة بالنطاق السلوكي وليس بصيغة الاختبار المحدد.

العيوب تستخدم تقديرات المحكمين لمواصفات البنود لتقدير درجة القطع لمفردات الاختبار ، بينما العلاقة بين المفردات ومواصفات البنود مباشرة ، لذا ليس من الضروري ان تكون تقديرات درجة القطع مماثلة لتقديرات مواصفات البنود (١٩٨٦، Berk).

ثانياً :الطرق التي تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية

اشار علام بأن الباحثين ادركوا قصوراً ملحوظاً في الطرق التحكيمية ومن ذلك القصور اعتمادها على تحكيم مفردات الاختبار دون النظر الى الاداء الفعلي للمختبرين مما يؤدي ذلك الى الحصول على معايير غير متميزة بدرجة عالية من الصدق .

واشار كل من سكون وجاليون وفيرارا بأن درجة القطع المحددة بأستخدام بعض من الطرق التحكيمية تؤدي الى رسوب اكثر من نصف المتقدمين للحصول على ترخيص مهنة الطب في امريكا ، اما في حالة استخدام الطرق التجريبية فإنه يتبين خطأ تلك الدرجات حيث يجتاز معظم الراسبين الحد الادنى لمستوى النجاح (علام،١٩٩٥م ،٢٤١-٢٤٢) .

لذا اقترح علماء القياس العديد من الطرق المختلفة لتحديد مستويات الاداء التي تتضمن طرقاً تجريبية محاولين في ذلك تجاوز عيوب الطرق التحكيمية والحصول على مستويات اداء دقيقة يمكن الاعتماد عليها ومن هذه الطرق مايلي :

١-الطريقة التحكيمية المعززة بالمعلومات Informed Judgement method (popham,1981)

يتضح من مسمى هذه الطريقة أنها تعتمد على بيانات مسبقة عن اداء الطلاب حيث تفيد هذه البيانات في اعطاء دلائل للمحكمين عن المستوى العام للطلاب ، وبالتالي توضع هذه البيانات في اعتبار المحكمين لكي يسترشدوا بها اثناء تحديدهم للمعايير وفي هذا الشأن ذكر بيرك(١٩٨٦، Berk)، بأن هذه الطريقة تعتمد على بيانات تتعلق بالاداء الفعلي للافراد ، وتتضمن هذه البيانات احصاءات خاصة بالمفردات مثل معاملات الصعوبة والتمييز

ومتوسط درجات الاختبار بعد تطبيقه على العينة المناسبة من الافراد ، وبعض تعليقات واحكام الخبراء حول مفردات الاختبار ، واقتراحات المختصين مثل المدرسين والاداريين واساتذه الكليات واعضاء مجالس المدراس ، وتقدم جميع هذه المعلومات للمحكمين لفحصها وتبادل المقترحات حولها ومناقشتها واعادة النظر فيها وذلك لكي يتم التوصل الى اتفاق حول درجة القطع المناسبة .**المميزات** سهلة التطبيق وسهلة الفهم ويسهل حسابها ، كما انها تجمع وتوحد بيانات الاداء والبيانات المتميزة من مدى واسع بواسطة المجموعات المختصة بتلك البيانات (Berk ،١٩٨٦).

العيوب : من عيوب هذه الطريقة انها تتطلب جمع كثير من البيانات التجريبية والمعلومات المتعلقة بالاختبار ومحتواه من مصادر متعددة ، وقد يستغرق ذلك وقتاً وجهداً كبيرين ، ويبدو ان تحليلات المحكمين للبيانات غير منظمة ، كما تبدو الحاجة لوجود خطوط تفصيلية تسهل استخدام انواع البيانات المختلفة والمناقشة الجماعية للاداء الفردي ، والتي لا تبدو مفيدة في هذه الطريقة بسبب تأثيراتها المعيارية (Berk ،١٩٨٦).

٢-طريقة انجوف التجريبية المعدلة Modified Angoff empirical mclean&Halpin,1984

ذكر بيرك(Berk ،١٩٨٦) ، ان هذه الطريقة عبارة عن تعديل لطريقة انجوف التحكيمية وتتطلب هذه الطريقة المعدلة تزويد المحكمين ببعض البيانات المستمدة من التجريب الميداني لمفردات الاختبار ، وذلك لكي يسترشد بها المحكمين في عملية التحكيم ، ويطلب من كل محكم تقدير احتمال اجابة الافراد ذوي الحد الادنى من الكفاية المطلوبة لمفردات الاختبار ككل وليس من كل مفردة على حده كما في الطريقة السابقة (التحكيمية)،وبعد الحصول على

هذه التقديرات يتم تزويد المحكمين ببيانات تجريبية عن مفردات الاختبار تشتمل على متوسط صعوبة كل مفردة ، ثم يطلب منهم تعديل تقديراتهم الاحتمالية في ضوء تلك البيانات التجريبية ، ويكون الحد الادنى لمستوى الاجتياز في الاختبار هو متوسط مجموع تقديراتهم المعدلة .

المميزات

سهلة التطبيق، سهلة الفهم ويسهل حسابها وتعمل على ربط المعيار بالكفايات المقاسة ، وتدمج معامل صعوبة البند الحقيقي خلال عمليات التحكيم .

العيوب

ان عملية استخدام تقديرات المحكمين للكفايات يجب ان تكون على درجة عالية من التوافق والانسجام ، وذلك بين المفردات والكفاية ولا يمكن افتراض ان تكون التقديرات الواقعية متطابقة (١٩٨٦، Berk).

واضاف جايقر Jaeger بأنه اجريت تعديلات في عام ١٩٧٨ م على طريقة انجوف الاساسية وذلك بهدف تبسيط مهمة المحكمين بحيث يطلب من كل محكم تقييم كل مفردة من مفردات الاختبار بالاجابة ب (نعم) او(لا) ، ومن ثم يتم تحديد الحد الادنى للكفاية وبعد الحصول على هذه التقديرات يستعرض عدد مختلف من المحكمين ثلاث انواع من البيانات التجريبية ، وذلك خلال ثلاث لقاءات منفصلة وتشتمل هذه البيانات التجريبية على التوزيع التكراري لدرجات الطلاب في الاختبار والاساليب الاحصائية الوصفية ودرجات القطع المحددة من قبل المحكمين ، ثم يطلب من عينات المحكمين المختلفة تعديل تقديراتهم بناءً على تلك البيانات التجريبية ، وتصبح درجة القطع عبارة عن متوسط تلك التقديرات

التي حددتها عينات المحكمين في اللقاءات الثلاثة المنفصلة (Linn,1989).

المميزات تتيح للمحكمين فرصة لتحسين تقديراتهم بناءً على ثلاث انواع من البيانات المختلفة ، ويحدد درجة القطع عينات مختلفة من المحكمين وتميل هذه الطريقة الى زيادة ثبات درجة القطع (١٩٨٦، Berk).

٣-طريقة توفق بين الطرق المطلقة والطرق النسبية Absolute Relative Comromise(I)(II) اشار بيرك (Berk,1986) ، الى ان هذه الطريقة تعمل على التوفيق بين الاحكام المحكمين المتعلقة بإداء الافراد في الامتحان والاداء الفعلي المستمد من البيانات التدريبية المبينية من تطبيق الاختبار على العينة ، وفي هذه الطريقة يطلب من المحكمين تحديث قيمتين هما :

أ-الحد الادنى للنسبة المئوية لعدد المفردات التي يجب ان يجيب عليها الفرد اجابة صحيحة لكي يعتبر ناجحا في الاختبار ، ويرمز لهذه النسبة بالرمز (K) .

ب-النسبة المئوية لعدد الافراد المتوقع اجتيازهم للاختبار ويرمز لهذه النسبة بالرمز (V)، ويمكن التوصل الى معادلة تربط بين النسبتين (K)، (V)، والتي يمكن استخدامها لتعديل تلك القيم التي لا تتفق مع التوزيع الفعلي لدرجات الاختبار ، اذ يمكن رسم منحني للمتوسطات والانحرفات المعيارية لقيم (K)، (V)، التي حددها المحكمون ، وكذلك النسبة المئوية لعدد الافراد المتوقع نجاحهم ويرمز لها بالرمز (y) كدالة تناقصية لدرجات الاختبار ويرمز لها بالرمز(x)، ثم يرسم منحني لتحديد النقطة ويرمز لها بالرمز (I) والتي تمثل الحد الادنى لمستوى النجاح ،بحيث توفق بشكل امثل

بين المستويات المطلقة(التحكيم) والمستويات النسبية (الاداء الفعلي للافراد في الاختبار) .

المميزات من مميزات هذه الطريقة انها توفق بين الاحكام المتعلقة بمفردات الاختبار وبين الاداء الفعلي للمختبرين اي بين المستويات النسبية . من مميزات هذه الطريقة انها توفق بين الاحكام المتعلقة بمفردات الاختبار وبين الاداء الفعلي للمختبرين اي بين المستويات المطلقة والمستويات النسبية كما تضع في اعتبارها الحد الاعلى للنسبة المئوية للمفردات التي من الممكن ان يجيب عليها الفرد اجابة صحيحة (Kmax)،كما تقوم بتبديل النسبة المئوية من الافراد المتوقع اجتيازهم للاختبار (V) بأعلى نسبة مئوية مقبولة من الراسبين (Fmax)وبأدنى نسبة مئوية مقبولة من الراسبين (Fmin)، (Berk,1986) .

العيوب

من عيوبها انها تتطلب من المحكمين ان يوفرو تقديرين وهذا يجعل تنفيذ هذه الطريقة اكثر صعوبة من غيرها سواء في عملية التحكيم او من الناحية الاحصائية، وكذلك تتطلب جهدا كبيرا من جانب المحكمين ، بالاضافة الى ان تفسير المعيار من قبل المتخصصين قد يكون صعب جدا .

لذلك اجرى هوفستي تعديلات لهذه الطريقة حيث يتطلب من المحكمين تحديد اربع قيم :

أ-اعلى نسبة مئوية لعدد المفردات التي يمكن ان يجيب عليها الفرد اجابة صحيحة ويرمز لها بالرمز (Kmax) وهي عبارة عن درجة القطع المرضية حتى اذا حصل عليها جميع الطلاب .

ب–ادنى نسبة مئوية من عدد المفردات التي يمكن ان يجيب عليها الفرد اجابة صحيحة ويرمز لها بالرمز (Kmin) وهي عبارة عن درجة القطع الدنيا التي لايمكن لاي طالب ان يحرز اقل منها .

ج–اعلى نسبة مئوية مقبولة من الراسبين ويرمز لها بالرمز (Fmax).

د–ادنى نسبة مئوية مقبولة من الراسبين ويرمز لها بالرمز (Fmin).

وبأستخدام توزيع درجات الاختبار المتجمع (تراكمي) فأننا نستطيع تحديد العلاقة التجريبية بين (K)و(F) ، ويكون المعيار هو نقطة التقاطع بين النموذج (المطلق)والمنحني (النسبي) .

ومن عيوب هذه الطريقة انها تتطلب جهدا كبيرا من المحكمين ، وتتطلب توفير اربعة تقديرات وهذا يجعل تنفيذ هذه الطريقة اكثر صعوبة من غيرها سواءً في عملية التحكيم او الاحصاء (Berk,1986) .

ثالثاً : طرق تعتمد على البيانات التجريبية وتسترشد بالتحكيم Empirical Judgmental

١–طريقة المجموعات المحكية Criterion groups (Berk,1984) .

اشار كل من بيرك وهامبلتون وآخرون (Berk,1986;Hambleton et al,1978) الى ان هذه الطريقة تتطلب انتقاء مجموعتين من الطلاب الاولى متقنة (تلقت التعليم) والثانية غير متقنة (لم تتلق التعليم)، ويكون مدى توزيع درجات الطلاب من الدرجة (صفر) الى الدرجة(ن) حيث (ن) تمثل عدد الاسئلة التي تقيس كل هدف على حدة ،وتتطلب الطريقة تقسيم الطلاب بناءا على درجة القطع (س) الى قسمين (متقنين –غير

متقنين)، وبافتراض ان الطلاب في المجموعة المتقنة (التي تلقت التعليم) عبارة عن طلاب متقنين فأنه يمكن تقسيمهم بناءعلى درجة القطع الى فئتين الاولى متقنين بالفعل (TM) والثانية غير متقنين غير حقيقين(FN)، وبالمثل يتم تصنيف طلاب المجموعة الثانية غير المتقنة (التي لم تتلق التعليم) الى فئتين الاولى غير متقنين فعلا(TM) والثانية غير متقنين غير حقيقين(FM) .

وتعتبر توزيعات درجات الطلاب في المجموعتين المتقنة وغير المتقنة المحددات الاولية (Brimary Determinants) لمدى دقة تصنيف الاختبار للطلاب كمتقنين وغير متقنين فعلاً للهدف المراد قياسه.

وتحدد قيمة درجة القطع(س) من خلال فحص خلية الاحتمالات لكل درجة قطع محتملة ، وتكون درجة القطع المثلى تلك التي تزيد نسبة التصنيفات الصحيحة (TM–TN).

ويتم تحديد درجة الاتقان بشكل اولي من خلال توزيعات تكرارات درجات الطلاب بمجموعتين المتقنة (التي تلقت التعليم) وغير المتقنة (التي لم تلتق التعليم حيث يصنف الطلاب بناءً على نقطة تقاطع توزيعين الى فئتين متقنين بالفعل وغير متقنين بالفعل للهدف المقاس .

٢–طريقة المجموعات المتضادة

LIVINGSTON&ZIEKY,1982 CONTRASTING GROUPS

ذكر كل من بيرك وجايقر بان هذه الطريقة مماثلة لحد ما لطريقة المجموعات المحكية وطريقة المجموعة الحدية ،ولكنها تختلف لكيفية تحديد درجة القطع في الاختبار ،ففي هذه الطريقة يقوم

المحكمون المتخصصون بتحديد مجموعتين من الطلاب بحيث تكون الاولى متقنة والاخرى غير متقنة بناء على مايلي:
المعلومات المتوفرة حول مستويات مهارات الطلاب .
استنادا على درجات الطلاب وادائهم داخل الفصل حيث يتم تصنيف الطلاب الى فئتين الحاصلون على تقدير (جيدجداً فما فوق)والثانية الطلاب الحاصلون على أعلى تقدير (مقبول فما دون)

ثم يطبق الاختبار على كلتا المجموعتين ويرسم منحنيات يمثلان التوزيعين التكرارين لدرجات الطلاب في الاختبار ،وتمثل نقطة تقاطع المنحنين درجة قطع الاختبار هذا الاقتراح يفترض تداخل التوزيعين التكرارين وعدم تطابقهما ،ويمكن تعديل هذه الدرجة اخذةً في اعتبارها خطأي التصنفين ، اي يمكن اختيار الدرجة التي تقلل من خطأ التصنيف الموجب (FALSE POSTIVE) او خطأ التصنيف السالب (FALSE NEGATIVE) .

العيوب

لها نفس عيوب طريقة المجموعات المحكية ، بالاضافة الى انه من الصعب تجاوز الخطأ الناجم من المحك (Linn,1989).

٣- طريقة المجموعة الحدية LIVINGSTON&ZIEKY,1982
BORDERLINE GROUPS

ذكر كل من بيرك وجايقر (Berk,1986;Jaeger,1989) بأن هذه الطريقة تتطلب من المحكمين انتقاء عينة من الطلاب الذين يرون بأن كفايتهم المراد قياسها تقع عند الحد الفاصل بين التمكن وعدم التمكن ، ومن المحتمل ان تكون نسبة الطلاب قليلة بالنسبة للمجموعة الكلية ،ثم يطبق الاختبار على افراد العينة وبعد ذلك يتم ايجاد وسيط درجات هذه المجموعة ليكون بمثابة مستوى الاداء المطلوب .

المميزات

تمتاز هذه الطريقة بأنها سهلة الفهم والتطبيق والاحصاء غير أنها تحتاج الى عدد كاف من الطلاب وذلك لكي نحصل على تقديرات مستقرة نسبياً .

العيوب

من عيوبها انه يصعب على المحكمين او المدرسين انتقاء المجموعة الحدية ، اذ ان تعريف او تحديد الطالب الذي يقع عند الحد الفاصل بين الاتقان اصعب بكثير من تحديد الفرد المتقن او غير المتقن ، كما تتدخل بعض المتغيرات في عملية انتقاء افراد هذه المجموعة وقد يصعب تحديد الحجم المناسب للعينة .

٤-طريقة المحك معياري المرجع NORM REFENCEED CRITERION

ذكر بيرك (Berk,1986) بأنه في هذه الطريقة يتم اختيار المعيار بناءً على العلاقة بين درجات القطع في الاختبار محكي المرجع وبين درجات الاداء في الاختبار معياري المرجع (NRT) ويتم اختيار درجة القطع اولاً في الاختبار معياري المرجع بطريقة اختبارية (على سبيل المثال: مناظرة الدرجة الخام بالدرجة المئينية ٥٠ او ٧٥) ثم تختار درجة القطع في الاختبار محكي المرجع على اساس انها تزيد مايلي :

أ-نسبة ثبات القرارات الناتجة عن تطبيق الاختبار معياري المرجع ويرمز لها بالرمز (p) .

ب-نسبة ثبات القرارات المصححة من عامل الصدفة ويرمز لها بالرمز (K) .

ج-تعيين دالة الفقد (Loss Function) ويرمز لها بالرمز(L).

د-سهلة التطبيق والفهم والاحصاء ويمكن توفير بيانات محك الاختبار معياري المرجع كمحك يعد خطأ فني وغير منطقي .

٥-طريقة النتائج التربوية Eductional Consequences (Block,1972)

ذكر بيرك (Berk,1986) بأن هذه الطريقة تعتمد أساساً على اختيار المعيار من خلال العلاقة بين درجات الاختبار ومحك التعلم المستقبلي (نتائج انفعالية او معرفية) حيث يتوقع ان يزداد المنحني طرديا ليربط بين المتغيرين ، وتحدد درجة القطع وفقاً على اثرها في زيادة اداء المحك .

المميزات

سهلة التطبيق والفهم والحساب .

العيوب

ان معيار (التعلم المستقبلي) يحتاج الى تحديد واضح ولا تبدو الافترضات في هذا النموذج متطابقة مع البيانات الواقعية ، بالاضافة الى ان العلاقة بين المتغيرات المقترحة لاتوفر "دالة درجية "(Step Function) ، او خاصية فاعلية في المحك لتحديد درجة القطع .

٦-طريقة ابمريك لتقييم اختبار التمكن Emrick's mastery Testing Evaluation

ذكر ايمربك (١٩٧١) EMRICK بأن هذه الطريقة تناسب الاختبارات الفرعية القصيرة التي تتكون من خمس فقرات او اقل ، وتعتمد هذه الطريقة على نظرية القرار ، اذ يتم من خلالها تحليل ثلاثة من أخطاء القرار وهي الخطأ الاول (α) والخطأ الثاني (β)ونسبة الفاقد لأخطاء القرارين الاول والثاني(RR) لذلك يجب تحديد الخطأ الاول والخطأ الثاني ونسبة الفاقد وطول الاختبار،ويتم

حساب نسبة الفاقد من خلال تقييم الانواع المختلفة لاخطاء القرار ، ومن ثم تسجيل مجموع المخاطرات بتحديد درجة القطع باستخدام المعادلة التالية :

$$K = \frac{\text{Log } \frac{\beta}{1-\alpha} + 1/n(\log RR)}{\text{Log } \frac{\alpha\beta}{(1-\alpha)(1-\beta)}}$$

حيث ان :

K ترمز الى درجة القطع ،**α** ترمز الى الخطأ الاول.

B ترمز الى الخطأ الثاني ، **RR** ترمز الى نسبة الفاقد للخطأين الأول والثاني .

n ترمز الى عدد الاسئلة .

٧-طريقة نوفيك ومعاونيه The work of Novick & Collaborators

ذكر (١٩٧٦) Meskauskas بأن هذه الطريقة عبارة عن مزيج من الطرق المختلفة ،وتهدف الى تحديد كل من درجة القطع والتوزيع الاحتمالي للاداء والفواقد الناتجة عن الخطأ الاول (α)والخطأ الثاني (β)ويفيد تحديد هذه المتغيرات في زيادة دقة القرارات المتخذة ، حيث يتم وضعها في شكل توزيع احتمالي ، ويضاف الى هذا التوزيع درجات الاختبار بحيث ينتج عنه توزيع احتمالي اكثر حساسية ، ويتم تحديد القيمة المقدرة للمستوى الحقيقي للطالب المتقن عن طريق تحديد عينات من المفردات ، ثم يضرب الاحتمال الذي تزيد فيه الدرجة الملاحظة على التحصيل المعطى ، ثم تقارن النتيجة بالدالة المشابهة التي تشير الى الاحتمال الذي لاتزيد فيه الدرجة الملاحظة .

٨-طريقة روبابش للدرجة الحقيقة الثنائية Roudabush's Dicotomous True Score

ذكر (١٩٧٦) Meskauskas بأن طريقة روبابش تستخدم مع الاختبارات ذات الفقرة الواحدة ، وتشمل هذه الطريقة على نموذجين من النماذج القطعية وكلا النموذجين يشتملان على مقياس ثنائي للدرجة الحقيقة ، وتقوم هذه الطريقة على تحديد اربعة من الاخطاء المحتملة وهي α_2،α_1(احتمال ان يكون الفرد المتقن ضمن غير المتقنين في المحك) و β_2، β_1 (احتمال ان يكون الفرد المتقن ضمن غير المتقنين في المحك) ويتم تحديد الاخطاء الاربعة وفق العلاقات التالية :

$$\alpha_2=P(X \geq X^o \backslash T=0) \quad \alpha_1=P(X \geq X^o \backslash T=0)$$
$$\beta_2=P(X \geq X^o \backslash T=1) \quad \beta_1=P(X \geq X^o \backslash T=1)$$

حيث ان :

T = الدرجة الحقيقة X = الدرجة الناتجة XC = درجة القطع

ونستخدم العلاقات التالية لتحديد درجة القطع وهي كالتالي :

$$F_{00}=N_0(1-a_1)(1-a_2)+N_1\mathcal{B}_1\mathcal{B}_2$$
$$F_{01}=N_0(1-a_1)a_2+N_1\mathcal{B}_1(1-\mathcal{B}_2)$$
$$F_{10}=N_0 a_1 (1-a_2)+N_1(1-\mathcal{B}_1)\mathcal{B}_2$$
$$F_{11}=N_0 a_1 a_2+ N_1(1-\mathcal{B}_1)(1-\mathcal{B}_2)$$

٩-طريقة ملمان ذات القرار المبني على الحدين MILLMAN'S BINOMIAL BASED DECISION

ذكر كل من هامبلتون ، وآخرون وميسكاوسكاس (Hambleton et al ,1978; Meskauskas ,1976) بأن ملمان milman طوّر مجموعة من الجدول التي تعتمد على التوزيع ذو الحدين ، والذي يشير الى كل من تقديرات الخطأ المتوقعة لمجموعة متعددة من الدرجات الحقيقة وطول الاختبار ودرجة القطع ، وهذه الطريقة

تقوم على افتراض ان الاختبار يتكون من مجموعة عشوائية من الدرجات تتروح مابين (١-صفر) ، ويستخدم التوزيع ذو الحدين لايجاد البيانات حول قيم محددة مفترضة للدرجة الحقيقة في كل من أطوال الاختبار والاختبارات الكبيرة والتوزيع الهندسي الفوقي او ذي الحدين ، فيتم التنبؤ بالتكرارات المتصل، ثم يرسم التكرار المتصل للنسب المئوية المختلفة من الاسئلة التي تمت الاجابة عليها بشكل صحيح .

١٠-طريقة ديفز ودايموند (١٩٧٤) THE DAVIS DIAMOND BYESIAN METHOD

اشار ميسكاوسكس (Meskauskas ، 1976) الى ان هذه الطريقة تعتمد على نموذج بييز ، وقد حدد كل من ديفز ودايموند جدولاً يضع تقديرات تعتمد على الدرجة الملاحظة واحتمال ان يكون بالمستوى الحقيقي لكفاءة الطالب بالقرب او أعلى من المستويات المختارة .

بالنسبة للاختبار (ذو خمس فقرات) فأن احتمال ان يكون الطالب قد احرز خمس درجات وله مستوى كفاءة حقيقي يساوي (٠,٩٩) او أعلى من ذلك يساوي (٠,١١٣٥) وهو رقم أدنى ، ويعرض الجدول درجات القطع للاختبارات حتى عشرون فقرة .

فاذا اراد صانع القرار ان يقتطع نسبة ٨٥% من التصنيفات الحقيقة للطلاب مع وجودمستويات كفاية عند نسبة ٩٠% او اعلى فيجب ان تكون هنالك مجموعة متساوية لعدد المفردات في الاختبار ، ولتحقيق هذا الهدف يجب استخدام اختبار تتروح فقراته مابين ١٢ الى ٢٠ فقرة .

أسس تقويم طرق تحديد درجة القطع

ذكر بيرك (Berk,1986) بأنه يوجد معيارين لتقويم طرق تحديد درجة القطع وهما كما يلي :

١- الكفاية الفنية Technical –Adequacy.

أ)ان تحدث الطريقة تصنيفاً مناسباً للمعلومات Appropriate Classification of information

بمعنى ان تحدث الطريقة تصنيفات ثنائية مثل الاتقان /عدم الاتقان ، الفشل /النجاح ، ويعتبر التصنيف اساسياً لمعظم القرارات المتخذة.

ب)ان تكون حساسة لاداء المختبر Sensitive Examinee Performance

يجب ان تكون الطريقة حساسة لمستويات صعوبة الاسئلة المختلفة ولأي خصائص اخرى مثل معامل التمييز .

ج)ان تكون الطريقة حساسة لكل من المتعلم او لعملية التدرب Sensitive to Instruction or Training

يجب ان تراعي الطريقة عملية التعلم والتدريب الفعلي الذي تلقاه الطالب بحيث تقاس المهارات والكفايات التي تلقاها الطالب بالفعل .

د)ان تتضمن الطريقة طرقاً احصائية مناسبة Statistically Sound

يجب استخدام الطرق الاحصائية المناسبة ، ويجب ان تفسر بشكل صحيح .

هـ)ان تحدد الطريقة المعيار الصحيح Identify the standard

بمعنى انه يجب ان تراعي الطريقة اخطاء القياس.

و)ان تثبت الطريقة صدق القرار Decision validity evidence

بمعنى انه يجب الحصول من الطريقة تقدير احتمالات قرارات التصنيفات الصحيحة والخاطئة .

٢-الجانب التطبيقي (العملي) Practicability

أ)ان تكون الطريقة سهلة التطبيق Easy to implement

يجب ان تكون الخطوات المستخدمة في تطبيق الطريقة واضحة وسهلة الفهم للمشاركين فيها ، وان تتم في قدر معقول من الوقت .

ب)ان تحسب الطريقة بسهولة Easy to Compute

يجب ان تكون الطرق الاحصائية المستخدمة للحصول على المعيار النهائي سهلة الاستخدام سواء باستخدام الالة الحاسبة او البرامج الاحصائية المتوفرة في الحاسب الآلي الشخصي او المركزي .

ج)ان قابلة للتفسير للعامة Easy to interpret to lay people

يجب ان تكون الطريقة المستخدمة في تحديد درجة القطع قابلة للتفسير وقابلة للفهم ، ويجب ان تكون تفسيرات الطريقة واضحة وسلة الادراك والفهم للعامة والمهتمين على حد سواء .

د)ان تكون الطريقة جديرة بثقة العامة Credible to lay people

يجب ان تكون الطريقة مقنعة وجديرة بثقة العامة ،كما اضاف بيرك(Berk,1986) بأن هناك عدة عوامل تساعد على اختيار الطريقة المناسبة وهي:

١-مدى اهمية القرارات المبنية على تحديد درجة القطع .

٢-الكمية المتاحة من الوقت لتحديد درجة القطع .

٣-مدى توفر المصادر (المادية -الطاقة البشرية)لانجاز مهمة تحديد درجة القطع .

كفاية المحكمين (بعض الطرق تتطلب دراية اكثر بالمحتوى المراد قياسه وبمستوى الطلاب التحصيلي).

٤-مدى مناسبة طريقة تحديد درجة القطع لنوعية الاختبار المراد قياسه .

ثم استخدم بيرك (Berk,1986)طريقة بديلة تهدف لايجاد معامل الصدق لكل درجة قطع محتملة (س)، لذا فإن متغيرات التصنيف تعطي قيم ثنائية هما :

أ)متغير التصنيف المنبئ The predictor Classification

حيث يقدر للطالب القيمة (١)عندما تكون درجته أعلى من درجة القطع ، ويقدر للطالب القيمة (صفر)عندما تكون درجته أدنى من درجة القطع .

ب)متغير التصنيف المحك Criterion Classification

يقدر لكل طالب ضمن مجموعة المتقنين القيمة (١)بينما يقدر لكل طالب ضمن مجموعة غير المتقنين (صفر)،ويحسب معامل الصدق من قيمة معامل (فاي) بين المتغيرين الثنائيين.

يلي ذلك حساب المنفعة والضرر النسبيين لاخطاء التصنيف ،ويمكن تمثيل قيمة الضرر للخطأ(FN) بالقيمة (-١) وللخطأ (FM) بالقيمة(-٢) ، في المقابل فإنه يتم حساب الفائدة النسبية لتحديد الاتقان بالفعل (TM)بالقيمة (٢) ولتحديد عدم الاتقان بالفعل (TN) بالقيمة (١) ، ويتم حساب الاحتمالات لاي اختبار تتبع الخطوات الاتية :

١-حساب نتائج الاحتمالات الاربعة التالية :

-طلاب متقنون فعلاً True Master.

-طلاب متقنون غير حقيقين False Master .

-طلاب غير متقنين فعلاً True Non Master.

--طلاب غير متقنون غير حقيقين False Non Master .

وتم ذلك بتقسيم مجموع عدد افراد العينة (٢١٩) طالباً على تكرار كل احتمال من الاحتمالات الاربعة .

ب: تم حساب نسبة احتمالات القرارات الصحيحة وفق الصيغة التالية :

القرارات الصحيحة =احتمال (طلاب متقنين حقيقيين +احتمال (طلاب غير متقنين حقيقيين)

P(TN) P(TM)

ج)تم حساب نسبة احتمالات القرارات الخاطئة وفق الصيغة التالية :

القرارات الخاطئة =احتمال (طلاب غير متقنين غير حقيقيين) +احتمال (طلاب متقنين غير حقيقيين)

P(FN) P(FM)

د)تم حساب معامل صدق درجات القطع المحتملة لقياس مدى تقدير التصنيف المتنبأ به (بناءً على درجة القطع المحتملة)ويتم ذلك وفق الصيغة التالية :

$$\phi = \frac{P(TM) - BR(SR)}{BR(1- BR)SR(1-SR)}$$

حيث ان :

BR(Base Rate)=احتمال الاتقان في المجموعة .

1- BR = احتمال عدم الاتقان في المجموعة .

SR= احتمال المقنيين المتنبأ بهم في المجموعة .

1-SR= احتمال غير المقنيين المتنبأ بهم.

وتم حساب الاحتماليين SR-BR وفق الصيغة التالية:

BR=P(FN)+P(TM)

SR=P(TN)+P(FM)

وتعتبر القيمة (+١) اعلى درجة لمعامل الصدق .

خامساً: تم حساب المنفعة والضرر النسبين لاخطاء التصنيف .

ويحسب الضرر المتوقع (&) للقرار الواحد (K) وفق الصيغة التالية:

K= P(FN)(D1)+P(FM)(D2)

حيث ان:

D1 =- ٢ ، D2 =- ١

ثم قامت الباحثة بحساب الفائدة المتوقعة (u) للقرار الواحد (k) وفق الصيغة التالية :

u

K=P™(U-1)+P(TN)(u2)

حيث ان **u1=٢** ، **u2=١** ،

الفصل الثالث

منهجية واجراءات البحث

منهجية البحث

استخدمت الباحثة المنهج الوصفي المقارن لانها تسعى الى وصف طرق تحديد درجة القطع المختلفة واستخدمت منهج الدراسات المقارنة لانها ركزت على مقارنة درجات القطع للاختبار والناتجة عن تطبيق طرق تحديد درجة القطع في الاختبارات محكية المرجع وذلك في ضوء متغير طول الاختبار

مجتمع وعينة البحث يتمثل مجتمع البحث بطلبة قسم العلوم التربوية والنفسية في كلية التربية بالجامعة المستنصرية للعام الدراسي ٢٠١٠/٢٠١١ والبالغ عددهم (١٩٥) طالباً وطالبة وبواقع (٩٢) طالب و (١٠٣) طالبة ، وقد قامت الباحثة بأختيار جميع الطلبة ليمثلوا عينة البحث الحالي .

اداة البحث تتمثل اداة البحث بالصورة الاولية للاسئلة التي تقيس الاهداف السلوكية لمادة علم نفس الشخصية المقرر على طلبة المرحلة الثالثة في اقسام العلوم التربوية والنفسية بكلية التربية في الجامعة المستنصرية

وقد اتبعت الباحثة عند صياغتها للصورة الاولية للاسئلة الخطوات الاتية :

تحديد المحتوى المراد قياسه: تحديد (المجال السلوكي)الذي تقيسه الاسئلة وهذا يتضمن تحديد محتوى المقرر الدراسي المراد قياسه وتحليله الى عناصره الاساسية ، حيث انحصر مجال اهتمام الباحثة في محتوى مادة علم نفس الشخصية والذي يدرس بواقع (٣) ساعات اسبوعياً لطلبة قسم العلوم التربوية والنفسية في كلية التربية بالجامعة المستنصرية اذ قامت بتحليل محتوى هذه المادة وتحديد عناصرها الاساسية التي يتكون منها مما يساعد في تحديد الاهداف السلوكية لعناصر المقرروذلك بالرجوع الى ماهو مثبت في اللجنة القطاعية الخاصة بمقررات مناهج كليات التربية الخاصة بقسم العلوم التربوية والنفسية لسنة ٢٠٠١، والجدول (٢) يوضح ذلك.

جدول (٢)

العناصر الاساسية لمحتوى مادة علم نفس الشخصية

ت	عنوان الفصل
١	مفهوم الشخصية عبر التراث السيكولوجي
٢	ابعاد بناء و نمو الشخصية
٣	نظريات الشخصية التي تستند الى الحتمية التكوينية
٤	نظريات الشخصية التي تستند الى الحتمية البيئيةو نظريات الشخصية التي تستند الى الحتمية التفاعلية
٥	الشخصية و الصحة النفسية
٦	الشخصية بين السواء والمرض
٧	تقييم الشخصية
٨	علاج الشخصية

وقد قامت الباحثة بعرض هذه الفصول والمواضيع بشكل استبانة على التدريسين الذين يقومون بمهام التدريس في الاقسام الاختصاص في كليات

التربية بهدف تحديد الفصول التي يقومون بتدريسها فعليا خلال السنة الدراسية خصوصا انه لايوجد كتاب مقررلديهم ،بل انهم يتبعون نظام المحاضرات في القاء المعلومات والتفاصيل الخاصة بكل فصل من فصول المقرر بحسب اطلاعهم المتجدد على الموضوعات العلمية الخاصة بهذه المادة وقد تم حذف الفصلين السابع والثامن لانها لم تحظ بموافقة غالبية التدريسين للقيام بتدريسها للطلبة وقد تبين بالرجوع الى اراء الخبراء والمتخصصين انه يمكن الاعتماد على نسبة ٨٠٪فاكثر من اتفاق اراء تدريسي مادة علم نفس الشخصية كخطوة اولى لتحديد محتوى مادة علم نفس الشخصية.

تحديد مفردات الاختبار التحصيلي : ويتضمن الخطوات التالية :

اولاً اعداد جدول المواصفات : يتم اعداد هذا الجدول بعد الفحص الدقيق لمحتوى المادة الدراسية المراد قياسها وذلك بعد تحليل المحتوى الى عناصره الاساسية من ناحية وتحديد مستويات الاهداف السلوكية من ناحية اخرى ومن ثم تحديد الاهمية النسبية للموضوعات والاهداف ، وقد قامت الباحثة بأعداد جدول المواصفات والجدول (٣) يوضح ذلك .

جدول (٣)

" جدول المواصفات لمادة علم نفس الشخصية مبيناً فيه عدد الفقرات و الاوزان المئوية لكل فصل"

الاهداف / المحتوى	الاهمية النسبية للمحتوى	عدد الاهداف السلوكية	٦٠% تذكر	٢٠% فهم	٢٠% تطبيق	عدد فقرات الفصل
الفصل الاول	١١،٠٢%	١٤	٨	٣	٣	١٤
الفصل الثاني	١٤،٩٦%	١٩	١١	٤	٤	١٩
الفصل الثالث	٣،٦٢%2	٣٠	١٨	٦	٦	٣٠
الفصل الرابع	٢٣،٦٢ %	٣٠	١٨	٦	٦	٣٠
الفصل الخامس	459،%	١١	7	٢	٢	١١
الفصل السادس	٣3،%17	٢٤	١٦	٤	٤	٢٤
	١٠٠%	١٢٨	٧٨	٢٥	٢٥	١٢٨

وبعد ان قامت الباحثة باعداد جدول المواصفات قامت بتحديد عدد الاسئلة التي ترتبط بالاهداف في موضوعات المادة موضوع الدراسة الحالية جدول (٤) .

جدول (٤)

جدول مواصفات مادة علم نفس الشخصية

ت	موضوعات المادة	المستويات المعرفية للاهداف						المجموع	
		تذكر		فهم		تطبيق			
		العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة
١	الاول	٨	%٨,٤٠	٣	%٢,٨٠	٣	%٢,٨٢	١٤	%١١,٠٢
٢	الثاني	١١	%١١,٤٠	٤	%٣,٨٠	٤	%٣,٨٠	١٩	%١٤,٩٦
٣	الثالث	١٨	%١٨	٦	%٦	٦	%٦	٣٠	%٢٣,٦٢
٤	الرابع	١٨	%١٨	٦	%٦	٦	%٦	٣٠	%٢٣,٦٢
٥	الخامس	٧	%٧,٢٠	٢	%٢,٤٠	٢	%٢,٤٠	١١	%٩,٤٥
٦	السادس	١٦	%١٣,٢١	٤	%٤,٤٠	٤	%٤,٤٠	٢٤	%١ ٧,٣٣
المجموع		٧٨	%٧٦,٢١	٢٥	%٢٥,٤	٢٥	%٢٥,٤	١٢٨	%١٠٠

ثانياً : كتابة مفردات الاختبار : اذ قامت الباحثة بتحديد الاهداف السلوكية لعناصر مادة علم نفس الشخصية والتي تغطي اربعة مستويات من النشاط الفعلي ، حيث حدد صلاح علام (١٩٨٦) بعض الملاحظات الخاصة بكتابة المفردات وهي :

أ- **اختيار نوع المفردات المناسبة لقياس الاهداف** : مفردات الاختيار والصواب والخطأ ، والمزواجة تصلح جميعها لقياس الاهداف ذات المستويات المعرفية الدنيا مثل (التذكر، الفهم ، التطبيق) بينما تصلح مفردات المقال لقياس المستويات المعرفية العليا علماً بأن :

التذكر : قدرة الطالب على استرجاع المعلومات او يتعرف عليها .
الفهم : ان يعيد الطالب صياغة المعلومات او يقدم لها وصفاً بأستخدام الفاظ من عنده .
التطبيق : ان يطبق الطالب قاعدة أو مبدأ في حل مشكلة معينة . (الدمنهوري،٢٠٠٠، ٣) .

ب- **تحديد عدد مناسب من المفردات** : حيث يمثل ذلك متغيرا مهما يؤثر في ثبات الاختبار وقدرته على التطبيق فكلما زاد عدد مفردات الاختبار زاد معامل ثبات الاختبار وفي الوقت نفسه يؤدي طول الاختبار الى ملل المختبرين مما يؤثر على ادائهم في الاختبار (علام ، ١٩٨٦،٤٠)
وبالرجوع الى مفردات الاهداف العامة التي وضعتها اللجنة القطاعية للعلوم التربوية والنفسية في وزارة التعليم العالي والبحث العلمي قامت الباحثة بتحديد بعض الاهداف السلوكية في شكلها النهائي ووضعها في شكل قائمة (ملحق ١)

ثالثاً: صياغة الاسئلة التي تقيس الاهداف السلوكية : قامت الباحثة بأختيار الشكل المناسب لاسئلة الاختبار وهي الاختيار من المتعدد ، وبالرجوع الى جدول المواصفات السابق والاسترشاد به قامت الباحثة بصياغة اسئلة الاختبار بحيث تكون مطابقة للاهداف السلوكية التي تقيسها.
-وبعد ان قامت الباحثة بصياغة اسئلة الاختبار عرضت الصورة الاولية على مجموعة من المحكمين المتخصصين (ملحق ٢) في الميدان للحكم على مدى صحتها العلمية ومن اسلوب صياغتها ومدى تغطيتها للاهداف السلوكية من حيث مستوياتها المعرفية المختلفة بحيث تأكد السادة المتخصصين بما يلي :

* هل تم صياغة اسئلة الاختبار بشكل علمي واكاديمي مناسب ؟
*هل شملت اسئلة الاختبار الاهداف السلوكية بمستوياتها المعرفية المختلفة؟

وقد وصل العدد النهائي لاسئلة الاختبار الى (١٢٨) سؤال وجميعها من نوع الاختيار من متعدد وتشمل كل مفردة على (٤) خيارات او بدائل احدها الاجابة الصحيحة .

وقد قامت الباحثة بعرضها على مجموعة من الخبراء المتخصصين في القياس والتقويم وعلم نفس الشخصية لمعرفة مدى اتفاق الاول مع الهدف السلوكي او المجال الذي تقيسه وقد قامت بوضع جدول يقوم الخبير بوضع علامة (نعم ، لا، غير متأكد) امام كل سؤال وذلك بعد اطلاع المحكم على ملحق (٣) ، موضح به توزيع الاسئلة على الاهداف التي تقيسها وبطلب من السادة الخبراء بأعطاء (+١) للسؤال اذا كان متأكداً من أنه يناسب الهدف الموضوع لقياسه ، وان يعطي تقدير (صفر) للسؤال اذا كان غير متأكد من ذلك ، وان يعطي (−١) للسؤال اذا كان غير متأكد من ذلك ، وان يعطي (−١) للسؤال اذا كان متأكد من أنها غير مناسبة لقياس الهدف.

صدق الاختبار : استخدمت الباحثة الصدق الوصفي والذي يعتمد على اراء الخبراء وذلك من خلال تحديدهم التجانس والاتساق بين السؤال والهدف الذي يقيسه وذلك يتضح من مراجعة البطاقة الخاصة بملحق (٣) ،و الجدول(٥) يوضح ذلك

جدول (٥)

نسبة اتفاق الخبراء على تطابق السؤال مع الهدف

تسلسل الاسئلة في الاختبار الرئيسي	عدد مرات الاتفاق	عدد مرات الاختلاف	نسبة الاتفاق
١، ٢، ٣، ٩، ١٥، ٢٠، ٢٥، ٢٧، ٣٠، ٣٧، ٤٤، ٤٥، ٤٧، ٥٢، ٥٣، ٥٧، ٥٨، ٥٩، ٦٥، ٦٦، ٦٧، ٦٨، ٧٨، ٧٩، ١١٢، ١١٤، ١١٥، ١١٦، ١١٧، ١١٨	١٠	–	%١٠٠
٧، ٨، ١٠، ١٣، ١٦، ١٩، ٢٤، ٢٨، ٣٣، ٣٤، ٣٦، ٣٨، ٣٩، ٤٦، ٤٨، ٥١، ٥٤، ٥٥، ٥٦، ٦٠، ٧١، ٧٢، ٧٦، ٧٧،	٩	–	%٩٠
٤، ٥، ٦، ١٢، ١٤، ٢١، ٢٢، ٢٣، ٢٦، ٢٩، ٣١، ٣٢، ٣٥، ٤٠، ٤١، ٤٩، ٥٠، ٦١، ٦٢، ٦٣، ٦٤، ٦٩، ٧٠، ٧٤، ٧٥، ٨٠، ٨١، ٨٢، ٨٣، ٨٤، ٨٥، ٨٦، ٨٧، ٨٨، ٨٩، ٩٠، ٩١، ٩٢، ٩٣، ٩٤، ٩٥، ٩٦، ٩٧، ٩٨، ٩٩، ١٠٠، ١٠١، ١٠٢، ١٠٣، ١٠٤، ١٠٦، ١٢٣، ١٢٤، ١٢٥، ١٢٦، ١٢٧.	٨	–	%٨٠
١٧، ١٨، ٤٢، ٤٣، ٧٣، ١٠٥، ١٢٨، ١٣٠١١٣	٧	٣	%٧٠

ومن خلال استعراض نتائج الجدول السابق يتضح تطابق الاسئلة مع الاهداف السلوكية التي تقيسها بدرجة تتروح بين (٧٠%-١٠٠%) وقد تم حذف (٧) فقرات لانها لم تحصل على نسبة اتفاق بين الخبراء اذ كانت اقل من المعيار المعتمد وهو ٨٠% .

ثبات الاختبار : للتأكد من ثبات الاختبار قامت الباحثة بحساب معامل ثبات الفاكروبناج وذلك للاختبارات الفرعية والاختبار الكلي والجدول (٦) يوضح ذلك .

جدول (٦)

معاملات ثبات الاختبارات الفرعية والاختبار الكلي لمادة علم نفس الشخصية

الاختبار	عدد الاسئلة	معامل الثبات
١	41	٠60,
٢	71	62٠,
٣	28	٠,٧٠
٤	29	72٠,
٥	11	58٠,
٦	12	٨6٠,
الاختبار الكلي	١٢٠	87٠,

ويلاحظ ارتفاع معامل ثبات الاختبار الكلي الامر الذي يدل على اتسام الاختبار الكلي بالثبات ويلاحظ وجود علاقة طردية بين قيم معامل ثبات ألفا وطول الاختبارات الفرعية .

تطبيق الاختبار : من اجل تحقيق اهداف البحث الحالي قامت الباحثة بتطبيق الصيغة الاولية للاختبار (ملحق ٤)وبمساعدة تدريسي المادة في القسم في الفترة الاخيرة من الفصل الثاني للعام الدراسي ٢٠١٠ -٢٠١١ وتم تصحيح الاختبار بأعطاء (درجة واحدة) للاجابة الصحيحة و (صفر) للاجابة الخاطئة وتم تفريغ البيانات في الاستمارات المعدة والاستعانة ببرنامج SPSS لاجراء التحليلات الاحصائية المناسبة فكانت نتائج التحليلات الاحصائية للاختبار الكلي كما يلي :

١-**المتوسطات والانحراف المعياري :** لتحليل اجابات عينة الطلبة على الاختبار قامت الباحثة بحساب متوسط درجات الطلبة وانحرافها المعياري وذلك للاختبار الكلي وللاختبارات الفرعية والجدول (٧) يوضح ذلك .

جدول (٧)

نتائج التحليلات الاحصائية المبدئية من تطبيق الاختبار على افراد العينة

الاختبار	الكفايات	عدد الاسئلة	المتوسط الحسابي	الانحراف المعياري
الاول	مفهوم الشخصية	١٤	١٣,٧٦	٣,٧٤
الثاني	ابعاد بناء و نمو الشخصية	١٧	١٦,٣١	٢،٩٩
الثالث	نظريات الشخصية التي تستند الى الحتمية التكوينية	٢٨	٢٧,٠٩	٠,٨٩
الرابع	نظريات الشخصية التي تستند الى الحتمية البيئيةو نظريات الشخصية التي تستند الى الحتمية التفاعلية	٢٩	٢٧,٧٩	٥,٠٥
الخامس	الشخصية و الصحة النفسية	١١	١١،٥٦	٠،٣٨
السادس	الشخصية بين السواء والمرض	٢١	٢٥,٦٧	٤,١٧
الكلي	اختبار علم نفس الشخصية	١٢٠	٧٣,٢٤	١٥,١٢

وبملاحظة الجدول يتبين ارتفاع متوسط درجات الطلبة في الاختبار الرابع مقارنة بالاختبارات الفرعية الاخرى .

٢-تحديد طول الاختبار: يعتبر متغير طول الاختبار احد متغيرات البحث الامر الذي يتطلب تعيين اختبارات مختلفة في عدد اسئلتها لقياس اثر المتغير على تحديد درجة القطع المستخدمة في الدراسة .

ونظراً لطول الاختبار الكلي (١٢٠) سؤال وتفاديا لاجهاد الخبراء المشاركين في تحكيم اسئلة الاختبار ، قامت الباحثة بأنتقاء عينة قصدية من اسئلة الاختبار الذي سبق تطبيقه على عينة البحث ، وذلك بناءاً على المعايير الثلاثة التالية :

أ)مستويات معامل الصعوبة المختلفة (صعوبة جدار صعبة، صعبة لحد – متوسطة الصعوبة ، سهلة لحد ما – سهلة –سهلة جدا)، اذا تم استثناء الفقرات التي كانت معاملات صعوبتها خارج مديات الفئات السبعة المحددة (ملحق (٥))

ب)معامل ارتباط السؤال المصحح مع مجموع الارتباط الكلي للاسئلة (ملحق (٦)) ،ونتج عن ذلك حذف (٣٦) سؤال ممن لم يكن معامل ارتباطها ذو دلاله احصائية واختبار (٨٤) سؤال من الاختبار الكلي تمثل الاختبار بصورة مبدئية

جـــ) قيمة معامل الثبات (α) اذا استثنى السؤال من مجموع الاسئلة الكلي والجدول (٨) يوضح ذلك .

جدول (٨)

تحديد طول الاختبار وفقا لمستوى الصعوبة ومعامل الارتباط وقيمة معامل الثبات

التسلسل	تصنيف المفردة	مستوى الصعوبة	معامل الارتباط	عدد الاسئلة
١	صعبة جدا	٩٠-٨٠	٠,٥٨	٩
٢	صعبة	٧٠٧٩	٠,٦٤	١١
٣	صعبة لحد ما	٦٩-٦١	٠,٨٣	١٤
٤	متوسطة الصعوبة	٦٠-٥٠	٠,٧٢	١٣
٥	سهلة لحد ما	٤٠-٣٠	٠,٧٨	١٣
٦	سهلة	٢٩-٢٠	٠,٦٨	١٢
٧	سهلة جدا	١٩-١٠	٠,٦٩	١٢
	المجموع الكلي	٩٠-١٠	-----	٨٤

ثم حددت الباحثة طول اختبارات الدراسة الحالية الثلاث وهي اختبار طويل يتكون من (٤٢) سؤال ، واختبار متوسط يتكون من (٢٨) سؤال ، واختبار قصير يتكون من (١٤) سؤال وقد قامت الباحثة بأختيار افضل (٤٢) سؤال من اسئلة الاختبار بصورته المبدئية (٨٤) سؤال بناءاً على معاملات صعوبة المفردات وبذلك تمكنت الباحثة من تحديد اسئلة الاختبار بصورته النهائية .

وقد أكتفت الباحثة بتطبيق الاختبار الطويل (٤٢) سؤال بناءاً على مستوى صعوبة الاسئلة مفردات الاختبارين المتوسط (٢٨) سؤال والقصير (١٤)

سؤال وبذلك تمكنت الباحثة من تحديد مفردات اختبارات البحث الثلاثة بصورتها النهائية وهي :

١-**الاختبار الطويل** : يتكون من (٤٢) سؤال تمثل جميع مستويات الصعوبة بواقع (٦) اسئلة لكل مستوى من المستويات السبعة ، والجدول (٩) يوضح ذلك .

جدول (٩)

تسلسل الاسئلة المنتقاة في الاختبار الكلي للاختبار الطويل

تسلسل الاسئلة المنتقاة في الاختبار الكلي	تصنيف الاسئلة	مستوى الصعوبة	عدد الاسئلة
٦٧،٤٩،٤٤،٣٨،٣١،٤	صعبة جدا	٨٠-٩٠	٦
١،٣،٩٣،٨٧،٦٣،٤٢،٢٣	صعبة	٧٠-٧٩	٦
٩٦،٧٥،٥٥،٤٧،٢٥،٦	صعبة لحد ما	٦١-٦٩	٦
١٠١،١٠٢،٩٠،٤٠،٢٨،١٠	متوسطة الصعوبة	٥٠-٦٠	٦
٩١،٨٤،٥١،٤٦،٢٤،١١	سهلة لحد ما	٣٠-٤٠	٦
٨٥،٧٩،٦٤،٤٣،١٧،١٤	سهلة	٢٠-٢٩	٦
٨٣،٥٩،٥٠،٣٥،١٥،١	سهلة جدا	١٠-١٩	٦
	الاختبار الفرعي الطويل	١٠-٩٠	٤٢

٢-**الاختبار المتوسط:** يتكون من (٢٨) سؤال تمثل جميع مستويات الصعوبة بواقع (٤) اسئلة لكل مستوى من المستويات السبعة والجدول (١٠) يوضح **ذلك**

جدول (١٠)

تسلسل الاسئلة المنتقاة في الاختبار الكلي للاختبار المتوسط

تسلسل الاسئلة المنتقاة في الاختبار الكلي	تصنيف السؤال	مستوى الصعوبة	عدد الاسئلة
٦٧،٤٤،٣٨،٤	صعبة جدا	٨٠-٩٠	٤
١٠٣،٨٧،٦٣،٢٣	صعبة	٧٠-٧٩	٤
٧٥،٥٥،٤٧،٦	صعبة لحد ما	٦١-٦٩	٤
١٠١،٩٠،٤٠،١٠	متوسطة الصعوبة	٥٠-٦٠	٤
٩١،٨٤،٢٦،٢٤	سهلة لحد ما	٣٠-٤٠	٤
٨٥،٧٩،٦٤،١٧	سهلة	٢٠-٢٩	٤
٨٣،٥٠،٣٥،١٥	سهلة جدا	١٠-١٩	٤
	الاختبار الفرعي المتوسط	١٠-٩٠	٢٨

٣-**الاختبار القصير** : يتكون من (١٤) سؤال تمثل جميع مستويات الصعوبة بواقع سؤالين لكل مستوى من المستويات السبعة والجدول (١١) يوضح ذلك .

جدول (١١)

تسلسل الاسئلة المنتقاة في الاختبار الكلي للاختبار القصير

تسلسل الاسئلة المنتقاة في الاختبار الكلي	تصنيف السؤال	مستوى الصعوبة	عدد الاسئلة
٤٤،٤	صعبة جدا	٩٠-٨٠	٢
٦٣،٢٣	صعبة	٧٩-٧٠	٢
٧٥،٥٥	صعبة لحد ما	٦٩-٦١	٢
٩٠،٤٠	متوسطة الصعوبة	٦٠-٥٠	٢
٨٤،٢٤	سهلة لحد ما	٤٠-٣٠	٢
٧٩،١٧	سهلة	٢٩-٢٠	٢
٣٥،١٥	سهلة جدا	١٩-١٠	٢
	الاختبار الفرعي القصير	٩٠-١٠	١٤

ثبات الاختبارات الثلاث : للتأكد من ثبات الاختبارات قامت الباحثة بحساب معامل ثبات الفا كردبناخ وذلك للاختبارات الثلاثة وكانت قيم معامل الثبات كما في جدول (١٢) .

جدول (١٢)

قيم معاملات الثبات للاختبارات الثلاثة

الاختبار	عدد الفقرات	معامل الثبات
الطويل	**٤٢**	**٠,٨٥**
المتوسط	**٢٨**	**%٨٠**
القصير	**١٤**	**٠,٧٥**

وبملاحظة الجدول (١٢) يتبين لنا ان جميع معاملات الثبات للاختبارات مناسبة ويمكن الحكم من خلال قيمها على اتسام الاختبارات بالثبات ، ويلاحظ ايضا اختلاف قيم معاملات الثبات بأختلاف طول الاختبار ، حيث ترتفع قيمة معامل الثبات بأرتفاع عدد مفردات الاختبار وهذا ناتج طبيعي .

تحديد درجة القطع لاختبار علم نفس الشخصية

لقد قامت الباحثة بتحديد درجة القطع بأربع طرق وهي :

١- **. طريقة أنجوف Angoffs Method**

: اعتمدت الباحثة ما ذكر في الاطار النظري حول تصميم استمارة تشتمل على تعليمات وارقام اسئلة الاختبار وكتب امام كل سؤال حرف مناظر للاجابة الصحيحة للسؤال داخل المربع ، وميزان يشتمل على عشر فقرات متساوية تقريباً (صفر-٠,١٠) ،(١١-٠,٢٠) ،(٠,٢١-٠,٣٠)،(٠,٣١-٠,٤٠)،(٠,٤١-٠,٥٠)،(٥١_٠,٦٠)،(٠,٦١-٠,٧٠)،(٠,٧١-٠,٨٠)،(٠,٨١-٠,٩٠)،(٠,٩١-١,٠) ، ثم طلب من كل محكم تقدير احتمال ان يعرف طالب الصف الثالث بقسم العلوم التربوية والنفسية الذي يمتلك الحد الادنى من الكفاية المطلوبة اجابة المفردة دون ان يلجأ الى تخمين الاجابة ثم يضع دائرة حول الرقم الذي يمثل هذه القيمة على الميزان المعطى ، ثم تجمع هذه القيم الاحتمالية المقدرة لكل مفردة على ذلك ايجاد متوسط هذه القيم لجميع المحكمين ولجميع مفردات الاختبارات (ملحق ٧) .

٢-**طريقة نيدلسكي** : يتم تحديد تقديرات المحكمين بعد فحص كل محكم اسئلة الاختبار ثم طلب منهم تصور مجموعة من طلبة الصف الثالث بقسم العلوم التربوية والنفسية لديهم الحد الادنى المقبول من الكفايات المتعلقة بمادة علم نفس الشخصية ، ثم طلب منهم تحديد البدائل الخاطئة التي من الممكن استبعاد اختيارها هؤلاء الطلبة لانها لا تمثل في نظرهم الاجابة الصحيحة للسؤال (ملحق ٨) وبذلك يكون الحد الادنى لمستوى اجتياز المفردة هو مقلوب عدد البدائل المتبقية للسؤال ثم تحققت الباحثة من ثبات تقديرات المحكمين وفق هذه الطريقة بحساب قيمة معامل سبيرمان براون اذ بلغت (٠,٦٢) وتعد هذه القيمة مقبولة وتم تحديد درجة القطع وفقا لهذه الطريقة وبحسب اختلاف طول الاختبار .

٣-**طريقة المجموعات المتضادة** : تم تحديد مجموعتين من الطلبة الاولى متقنة والثانية غير متقنة بناءاً على الاداء الفعلي (Actual data) لعينة الدراسة من الطلبة في مادة علم نفس الشخصية ، وتم تحديد محك يصنف الطلبة الى فئتين متقنة وغير متقنة لذا تم حساب معامل ارتباط بيرسون بين درجات جمع افراد العينة خلال الفصلين الدراسيين للعام الدراسي ٢٠١٠-٢٠١١ وبين درجاتهم في الكلية على الاختبار الطويل (٤٢) سؤال ، حيث يتبين ان معامل الارتباط بين درجات اختباري الفصلين (٧١) درجة وبين الاختبار الطويل (٤٢) سؤال الا على ارتباط (ر =٠,٦١) وبذلك استبعدت الباحثة درجات الفصلين لتدني معامل ارتباطهما ، ثم قامت بتحديد عينة قصد به تمثل المجموعة الاولى المتقنة بناءاً على درجاتهم المرتفعة (> او=٧٧) اي الحاصلين على نسبة (٩١,٦٧٪) فأعلى في الفصلين الدراسيين وبذلك اصبح عدد افراد عينة المتقنين (٣٢) طالبا وطالبة ، ثم حددت المجموعة الثانية غير المتقنة بناءا على درجاتهم المتدنية (< او =٢٨) اي الحاصلين على نسبة (٣٣,٣٣٪) فأدنى في اختبار الفصلين وبذلك اصبح عدد افراد عينة غير المتقنين (٣٥) طالبا

وطالبة ثم حصل على مجموع درجات المجموعتين المتقنة وغير المتقنة المحددتين وفق المحك السابق ذكره في الاختبارات الثلاثة ثم تم تمثيل توزيع مجموع تكرار الدرجات الكلية للمجموعتين بيانياً وبذلك تم تحديد درجات قطع الاختبارات الثلاثة من تقاطع المنحنيات

٤- **طريقة المجموعات المحكية** : قامت الباحثة بتحديد فئتين من الطلبة الاولى متقنة والثانية غير متقنة مستمداً في ذلك على الخطوات السابقة التي قام بها في المجموعات المتضادة مع اختلاف في درجة المحك (> أو = ٦٨) درجة اي الطلبة الحاصلين على نسبة٩٥ ٨٠٪ فأعلى ، (< أو = ٣٤) درجة اي الطلبة الحاصلين على نسبة(٤٨ ٤٠%) فأدنى ، وذلك بهدف زيادة عدد افراد العينة القصيدية ولتحديد درجة القطع للاختبارات المختلفة قامت الباحثة بإيجاد مجموع درجات المجموعتين المتقنة وغير المتقنة لجميع الاختبارات (طويل ، متوسط، قصير) بناءاً على درجات المحك ثم استخدام برنامج الحزمة الاحصائية للعلوم الاجتماعية لتمثيل توزيع مجموع تكرار درجات المجموعتين الكلية بيانيا وحصل على درجات القطع لجميع الاختبارات الثلاثة من نقاط تقاطع منحنيات التوزيعات التكرارية .

الوسائل الاحصائية :

استخدمت الباحثة الوسائل الاحصائية الاتية لتحليل بيانات البحث الحالي :

١-المتوسطات الحسابية والانحرافات المعيارية لوصف اداء الطلبة في الاختبار الرئيس (١٢٠) سؤال .

٢- معامل ثبات الفاكرونباح لتحديد معامل ثبات الاختبارات الفرعية والاختبار الرئيسي المطبق على عينة من الطلبة .

٣-معامل ارتباط المفردة المصحح مع مجموع الارتباط الكلي للمفردات لتحديد اسئلة الاختبار الطويل(٤٢) سؤال من اختبار علم نفس الشخصية الرئيس (١٢٠)

٤-معامل صعوبة المفردات اسئلة الاختبارات الثلاثة (الطويل ،المتوسط، القصير) من اسئلة اختبار علم نفس الشخصية الرئيس (١٢٠)

٥-تحديد درجة القطع وفق طريقة نيدلسكي بالصيغة المذكورة في ص (١٨)

٦-تحديد درجة القطع وفق طريقة المجموعات المحكية بالصيغة المذكورة في ص (٢٣)

٧-اسلوب تحليل التباين الاحادي لتحديد الفروق بين قيم درجة القطع للاختبارات الثلاثة (طويل ،متوسط، قصير) وفق (طريقة انجوف ونيدلسكي) .

الفصل الرابع

عرض النتائج ومناقشتها –التوصيات– المقترحات

سيتم عرض نتائج المتعلقة بهدف البحث المتمثل بتعرف اثر طول الاختبار على تحديد درجة القطع لاختبار علم نفس الشخصية على مستويين :

اولا:عرض ومناقشة نتائج اثر طول الاختبار على مستوى كل طريقة على حدة :

١-طريقة انجوف

تم تحديد درجة القطع وفق طريقة انجوف للاختبارات الثلاثة المختلفة (طويل ،متوسط، قصير) والجدول (١٣) يوضح ذلك :

جدول (١٣)

درجات القطع وفق طريقة انجوف بأختلاف طول الاختبار

طول الاختبار	عدد الاسئلة	درجة القطع	النسبة
طويل	٤٢	٣٥	٨٣,٣
متوسط	٢٨	٢٣,٥	٨٣,٩
قصير	١٤	١٢	٨٥,٧

وبملاحظة الجدول (١٣)يتضح وجود فروق صغيرة في نسب درجات قطع الاختبارات الثلاثة المحددة وفق طريقة انجوف ، وان درجة القطع تتناسب تناسباً عكسياً مع طول الاختبار اذ تقل قيمة درجة القطع كلما زاد عدد الاسئلة .

ولتحديد الفروق بين قيم درجة القطع للاختبارات الثلاثة استخدم تحليل التباين الاحادي للعينات غير المتساوية والجدول (١٤) يوضح ذلك .

جدول (١٤)

تحليل التباين الاحادي لقيم درجة القطع وفق طريقة انجوف بأختلاف طول الاختبار

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	القيمة الغائبة
بين الاختبارات	١١,٨٤٦	٢	٥,٩٢٣	٠,٩٠٥
داخل الاختبارات	٥٢٩,٩٠٢	٨١	٦,٥٤٢	
الكلي	٥٤١,٧٤٨	٨٣		

وتشير النتائج في الجدول اعلاه الى عدم وجود فروق دالة احصائياً عند مستوى دلالة (٠,٠٥) بين قيم درجات قطع الاختبارات الثلاثة وفق طريقة انجوف.

٢- طريقة نيدلسكي

تم تحديد درجة القطع وفق طريقة نيدلسكي للاختبارات الثلاثة المختلفة (طويل،متوسط،قصير) والجدول (١٥) يوضح ذلك .

جدول (١٥)

درجات القطع وفق طريقة نيدلسكي بأختلاف طول الاختبار

طول الاختبار	عدد الاسئلة	درجة القطع	النسبة
طويل	٤٢	٢٧	٦٤,٣
متوسط	٢٨	١٨	٦٤,٣
قصير	١٤	٩	٦٤,٣

وبملاحظة الجدول (١٥) يتضح تساوي نسب درجات قطع الاختبارات الثلاثة المختلفة وفق طريقة نيدلسكي (٦٤,٣) مما يدل على عدم وجود اثر لطول الاختبار على تحديد درجة القطع وفقا لطريقة نيدلسكي ، ولتحديد الفروق بين قيم درجة القطع وفق طريقة نيدلسكي بأختلاف طول الاختبار استخدم تحليل التباين الاحادي والجدول (١٦) يوضح ذلك

جدول (١٦)

تحليل التباين الاحادي لقيم درجة القطع وفق طريقة نيدلسكي بأختلاف طول الاختبار

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	القيمة الغائبة
بين الاختبارات	٦,٦٤	٢	٣,٣٢٠	٠,٧٨١
داخل الاختبارات	٣٤٤,٣٣١	٨١	٤,٢٥١	
الكلي	٣٥٠,٩٧١	٨٣		

وتشير النتائج في الجدول اعلاه الى عدم وجود فروق دالة احصائياً عند مستوى دلالة (٠,٠٥) بين متوسطات قيم درجات قطع الاختبارات الثلاثة الامر الذي يدلل على عدم وجود علاقة بين طول الاختبار ودرجة القطع .

٣-طريقة المجموعات المتضادة

تم تحديد درجة القطع وفق طريقة المجموعات المتضادة للاختبارات الثلاثة المختلفة (طويل،متوسط،قصير) والجدول (١٧) يوضح ذلك .

جدول (١٧)

درجات القطع وفق طريقة المجموعات المتضادة بأختلاف طول الاختبار

طول الاختبار	عدد الفقرات	درجة القطع	نسبة القطع
طويل	٤٢	٣٠,٥	٧٢,٦
متوسط	٢٨	١٨	٦٤,٣
قصير	١٤	٩,٥	٦٧,٩

وبملاحظة الجدول(١٧) يتضح اختلاف نسب درجات القطع وفق هذه الطريقة مما يدل على وجود اثر لمتغير طول الاختبار على

تحديد درجة القطع اذ حقق الاختبار الطويل اعلى نسبة (٧٢,٦) لدرجة القطع ، وفي حين حقق الاختبار المتوسط ادنى نسبة (٦٤,٣) لدرجة القطع .

٤-طريقة المجموعات المحكية

تم تحديد درجة القطع وفق طريقة المجموعات المحكية للاختبارات الثلاثة المختلفة (طويل، متوسط، قصير) والجدول (١٨) يوضح ذلك .

جدول (١٨)

درجات القطع وفق طريقة المجموعات المحكية بأختلاف طول الاختبار

طول الاختبار	عدد الفقرات	درجة القطع	نسبة القطع
طويل	٤٢	٢٧	٦٤,٣
متوسط	٢٨	١٨	٦٤,٣
قصير	١٤	٩	٦٤,٣

وبملاحظة الجدول(١٨) يتبين لنا ان نسب درجات القطع للاختبارات الثلاثة المختلفة متساوية وفق هذه الطريقة مما يدل على عدم وجود اثر لطول الاختبار على تحديد درجة القطع بهذه الطريقة مما يؤكد النتائج السابقة لكل من طريقة انجوف ونيدلسكي.

ثانيا ً : عرض ومناقشة نتائج اثر طول الاختبار على مستوى طرق الدراسة مجتمعة .

سيتم عرض ومناقشة اثر طول الاختبار بناءاً على نتائج طرق تحديد درجة القطع في الدراسة مجتمعة مما يساعد على مقارنة نتائج جميع طرق تحديد درجة القطع لكل اختبار من الاختبارات الثلاثة على حدة (طويل، متوسط، قصير) وهي كالتالي :

١- **الاختبار الطويل (٤٢) سؤال :** قامت الباحثة بأعداد جدول خاص بالاختبار الطويل (٤٢) فقرة اعتماداً على الجدوال

السابقة التي حددت فيها درجات القطع وصحت فيه عدد ونسبة الطلبة المتقنين وغير المتقنين والناتجة وفق الطرق المختلفة للاختبار الطويل (جدول ١٩) .

جدول(١٩)

تكرارات ونسب الطلبة الطلبة المتقنين وغير المتقنين بأختلاف الطرق للاختبار الطويل

الطريقة المستخدمة	درجة القطع		المتقنين		غير متقنين		العينة
	الدرجة	النسبة	العدد	النسبة	العدد	النسبة	
انجوف	٣٥	٨٣,٣	١٠	٥,١٣	١٨٥	٩٤,٨٧	١٩٥
نيدلسكاي	٢٧	٦٤,٣	٧١	٣٦,٤١	١٢٤	٦٣,٥٩	١٩٥
المجموعات المتضادة	٣٠,٥	٧٢,٦	٤٦	٢٣,٥٩	١٤٩	٧٦,٤١	١٩٥
المجموعات المحكية	٢٧	٦٤,٣	٧١	٣٦,٤١	١٢٤	٦٣,٥٩	١٩٥

وبملاحظة الجدول(١٩) يتبين لنا ان نسبة درجة القطع المحددة وفق طريقة انجوف كانت الاعلى (٨٣,٣) حيث نتج عنها ادنى نسبة للطلبة المتقنين (٥,١٣) يليها نسبة طريقة المجموعات المتضادة (٧٢,٦) التي نتج عنها ارتفاع ملحوظ في نسبة المتقنين (٢٣,٥٩) ، اما طريقتي نيدلسكي والمجموعات المحكية فقد تساوت النسبة فيها على الرغم من الاختلاف الكبير في اجراءات تحديدهما وترتب على هذه النسبة ارتفاع في نسبتي الطلاب المتقنين (٣٦,٤١) وانخفاض نسبة غير المتقنين (٦٣,٤١) مقارنة مع الطريقتين الاخريين .

ب-الاختبار المتوسط (٢٨) سؤال : قامت الباحثة بأعداد جدول خاص بالاختبار المتوسط (٢٨) سؤال اعتماداً على الجدول السابقة التي حددت فيها درجات القطع للاختبار المتوسط (٢٨) سؤال

وضحت فيه عدد ونسب الطلبة المتقنين وغير المتقنين والناتجة وفق الطرق المختلفة للاختبار المتوسط (جدول ٢٠)

جدول (٢٠)

تكرارات ونسب الطلبة المتقنين وغير المتقنين بأختلاف الطرق للاختبار المتوسط

الطريقة المستخدمة	درجة القطع		المتقنين		غير متقنين		العينة
	الدرجة	النسبة	العدد	النسبة	العدد	النسبة	
انجوف	٢٣,٥	٨٣,٩	١٤	٧,١٨	١٨١	٩٢,٨٢	١٩٥
نيدلسكاي	١٨	٦٤,٣	٧٠	٣٥,٩٠	١٢٥	٦٤,١٠	١٩٥
المجموعات المتضادة	١٨	٦٤,٣	٧٠	٣٥,٩٠	١٢٥	٦٤,١٠	١٩٥
المجموعات المحكية	١٨	٦٤,٣	٧٠	٣٥,٩٠	١٢٥	٦٤,١٠	١٩٥

وبملاحظة الجدول (٢٠) يتبين لنا تساوي نسبة درجة القطع للطرق الثلاث المختلفة نيدلسكاي، المجموعات المتضادة ، المجموعات المحكية) رغم الاختلاف الكبير في اجراءات تحديدها حيث نتج عن ذلك تساوي نسبة الطلبة المتقنين (٣٥,٩) وكذلك تساوي نسبة الطلبة غير المتقنين (٦٤,١) ، اما طريقة انجوف فقد حققت اعلى نسبة (٨٣,٩) درجة قطع حيث نتج عنها ادنى نسبة للطلبة المتقنين (٧,١٨) واعلى نسبة للطلبة غير المتقنين (٩٢,٨٢) .

ج- **الاختبار القصير (١٤) سؤال :** قامت الباحثة بأعداد جدول خاص بالاختبار القصير (١٤) سؤال اعتماداً على الجداول السابقة التي حددت فيها درجات القطع للاختبار القصير (١٤) سؤال وضحت فيه عدد ونسب الطلبة المتقنين وغير المتقنين الناتجة وفق الطرق المختلفة للاختبار القصير (جدول ٢١) يوضح ذلك .

جدول(٢١)

تكرار ونسب الطلبة المتقنين وغير المتقنين للاختبار القصير

الطريقة المستخدمة	درجة القطع		متقنين		غير متقنين		العينة
	الدرجة	النسبة	العدد	النسبة	العدد	النسبة	
انجوف	١٢	٨٥,٧	٧	٣,٥٩	١٨٨	٩٦,٤١	١٩٥
نيدلسكاي	٩	٦٤,٣	٧٦	٣٨,٩٧	١١٩	٦١,٠٣	١٩٥
المجموعات المتضادة	٩,٥	٦٧,٩	٥١	٢٦,١٥	١٤٤	٧٣,٨٥	١٩٥
المجموعات المحكية	٩	٦٤,٣	٧٦	٣٨,٩٧	١١٩	٦١,٠٣	١٩٥

وبملاحظة الجدول (٢١) تبين لنا ان طريقة انجوف حققت اعلى نسبة (٨٥,٧) ترتب عليها تدني مستوى الطلبة المتقنين وارتفاع نسبة الطلبة غير المتقنين ، تلت هذه الطريقة بفارق واضح طريقة المجموعات المتضادة (٦٧,٩) ترتب عليها ارتفاع في نسبة الطلبة المتقنين (٢٦,١٥) وتدني نسبة الطلبة غير المتقنين (٧٣,٨٥) ، اما بالنسبة لطريقتي نيدلسكاي والمجموعات المحكية فقد تساوت النسبة لدرجة القطع بينهما (٦٤,٣) حيث حققنا اعلى نسبة للطلبة المتقنين (٣٨,٩٧) وادنى نسبة للطلبة غير المتقنين (٦١,٠٣)

النتائج التي توصل اليها البحث:

١-تساوي نسب درجات القطع الناتجة عن تطبيق طريقتي نيدلسكاي والمجموعات المحكية بأختلاف طول الاختبار .

٢-ارتفاع نسب درجات القطع الناتجة عن تطبيق طريقة انجوف وذلك عند تطبيقها على الاختبارات المختلفة من حيث الطول (طويل، متوسط، قصير) .

٣-تقارب نسبتي الطلبة المتقنين True Method الى حد ما (٣٨,٩٧،٣٥,٩،٣٦,٤١) والطلبة غير المتقنين True Non Master (٦٤,١،٦١,٠٣،٦٣,٥٩) عند تطبيق طريقة المجموعات المحكية على الاختبارات المختلفة مقارنة بالطرق الاخرى .

٤- حققت طريقتي نيدلسكاي والمجموعات المحكية ادنى درجات قطع عند تطبيقهما على الاختبارات المختلفة (طويل،متوسط،قصير) .

التوصيات

١-تدريب المحكمين بصورة جيدة مما يساعد على الحصول على تقديرات دقيقة لدرجات قطع الاختبار .

٢-تحديد فئتي الطلبة المتقنين والطلبة غير المتقنين عند تطبيق طريقة المجموعات المتضادة بأعتماد درجات الطلبة في اختبارات تحصيلية سابقة في مادة الاختبار .

٣-تعريب او تصميم برامج خاصة تسهل عملية تحديد درجة القطع بالحاسب الالي .

المقترحات

١-اجراء دراسة لمعرفة اثر تدريب المحكمين على تحديد درجة القطع .

٢-اجراء دراسة لمعرفة اثر استخدام الطرق التي تعتمد على بيانات واقعية حول مستوى تحصيل الطالب .

٣-اجراء دراسة مماثلة على اداة الدراسة الحالششية تعتمد على متغيرات وعينة دراسية مختلفة

المصادر:

١) ابراهيم ، محمود ، ١٩٩٠ ، دراسة سيكومترية مقارنة لطرق حساب معامل ثبات الاختبارات المرجعة الى المحك ، رسالة ماجستير ، كلية التربية ، جامعة عين شمس.

٢)ابو حطب ، فؤاد (١٩٨٧ م) ، التقويم النفسي ، القاهرة، الانجلو المصرية .

٣)ابو علام،رجاء،١٩٨٧،قياس وتقويم التحصيل الدراسي،ط١،الكويت ،دار القلم.

٤)الاحمد ،أحمد (١٤١٣ هـ) ، تأثير طريقة اختبار فقرات الاختبار المحكي المرجع على خصائصة السيكومترية ، رسالة دكتوراه غير منشورة ، عمان ،الجامعة الاردنية .

٥)الشرقاوي ، انور وآخرون (١٩٩٦) ،اتجاهات معاصرة في القياس التقويم النفسي والتربوي ، القاهرة،مكتبة الانجلو المصرية .

٦)عبدالسلام، نادية (١٩٩١م) ،تعدد معطيات تحديد الدرجات الفاصلة في القياس محكي المرجع (دراسة تجريبية) حولية كلية البنات ، جامعة عين شمس ، المجلد ١، العدد١٦.

٧)_____ (١٩٩٢م) ، ثلاث محكات رئيسية لتحديد مستوى الاتقان في الاختبارات المحكية المرجع ، المجلة المصرية للدراسات النفسية ، العدد ٣.

٨)علام، صلاح الدين (١٩٨٥م) ،استخدام النموذج ذي الحدين في تقدير درجة القطع للاختبار محكي المرجع (دراسة احصائية) ، المجلة العربية للعلوم الانسانية ، جامعة الكويت .

٩)_____ (١٩٨٦م) ، تطورات معاصرة في القياس النفسي والتربوي ، الكويت :مطابع القبس التجارية .

١٠)_____ (١٩٩١م) ، دراسة مقارنة لبعض طرق تحديد مستويات الاداء في اختبار محكي المرجع ، المجلة المصرية للدراسات النفسية ، العدد ١: ٧٧-٩٦ .

١١)_____ (١٩٩٥م) ، الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية ، القاهرة : دار الفكر العربي .

١٢) Andrew،bj٠and hecht،j٠t٠،(1976)apreliminary investigation of two procedures for setting examination standard ،**educational and psychologyical measurement** 36:45-50 ٠

١٣)Angoff ,W.H.(1971),Scales,Norms and Equivalent Scores,In R.L.Thorndik(Ed.)**Educational Measurement** .Washington ,D.C:American council on education.

١٤)Berk,R.A(1980**).Criterion Referenced Measurement**. The State of Art.

15) ——— (1984)Aguide to criterion –referenced test construction baltimmmmore،london ،the johns hopppkins university press

16)_____ (1986). Consumers Guide To Setting Performance Standard on Criterion Referenced Test .**Review of Educational Research** .56:137-172.

17)Crocker.I.,& Algina ,J.(1986). **Introduction to classical & modern test theory** .New York Holt . Rinehart & Winston .

18)Cross ,I.H.,Impara,I.C.,Frary ,R.B.,And Jaeger , R.M.(1984).A comparison of Three Methods for Establishing Minimum Examination **,Jourunal of Educational Measurement. 21**:113-129 .

19)Ebel , r.l.(1979**).Essentials of Educational Measurement** Englewood cliffs .nj: prentice hall.

20)Emrick ,j.a(1971).an evaluation models for mastery testing ,**journal of Educational Measurement .8:**321-326.

21)Glassg٠a٠، (1970)standards and criteria،**journal of educational mearsurment**،vol٠15،no٠4،237-261٠

22)Glaser,r (1994). Instructional technology & the Measurement of learning outcomes :some questions :some questions .**Educational Measurement :issues &practice .13**:6-8.

23)Gray،w،m،،(1978): Acomparion of Piagetian theory&criterion referenced measurment،**review of educational research** ،48،p223-249،

24)Haldyna,T.M.,Roid,G.H.(1983): A comparation of two Approaches to Griterion –Raferen ced Test , Jourral of Educational Measurement**,** vol.20 ,No.3,pp271-281.

25)Haplin،g،،sigmon&haplin g،،(1983):mini competency standards set by three divergent groups of raters using three judgmental procedures implications for validity ،educational &**psychology ical measurement**،43،185-196،

26) Hambleton ,R.K.,&Novick,M.R.(1973).Toward An Integration of Theory and Method for criterion Referenced Tests. **JEM.**10:159-170.

27)_____ (1978), on the use of cut off score with criterion referenced test in instructional setting ,**jem** .**15**:277-290.

28)_____,Swaminathan ,H.,& Algina .(1976) Some contributions to the theory and practice of criterion referenced testing .In D.N.M. de gruijter, & van der kamp (Eds) ,Advances in psychological and **Educational Measurement** .New York : Wiley .

29) _____ & Swaminathan ,H, Algina. J & couison .D.B.(1978) .criterion referenced testing & measurement : A review of technical issues & development .**Review of educational research .48**:1-47.

30)_____ &Eignor ,D.R(1980) .Competency Test development Validation , and standard setting . INR.M.jeager &C.K.Tittle (Eds).**Minimum competency a achievement testing Motives ,models .measures and consequences .Berkeley ,CA:Mccutchan .**

31) Jaeger , R.M(1989).certification of student competence . IN linn. R.L.(Eds.)**educational measurement** .new York :macmillan .
32)Koffler , S.L.(1980) A comparison of Approaches for setting proficiency standards .jem.17:167-178.
33) Meskauskas ,J.A. (1976) .evaluation models for criterion referenced testing : views regarding mastery & standard setting .**Review of educational research .46**:133-158.
34)Nedelsky , L.(1954).Absolute grading standards for objective test **Educational and psychological Measurement .14:**3-19.
35)Nunnlly،j،c،(1978)**psychological theory**،2nd ed،،new York:mcgraw-hill،
36)Popham,W.J.(1978) **criterion referenced measurement** , englewood cliffs,NJ.prentic hall.
37) _____& Husek ,T.R.(1969) .Implications of criterion referenced measurement **.JEM.6**:19.
38)Sax &Newton ,J.W,(1997),principles of Educational and psychological Measurement and Evaluation ,4th ed. New York : Wadsworth publishing comany.
39)skakun،e،&kling ،s،(1980):comparability of methods for setting –jem 17،pp229-235،
40)Subkoviak ,M.J.(1980) .decision consistency approaches . In R.A. BERK(Eds.), **criterion referenced measurement** .the state of art PP.129-183.
41)Winer, B.J,(1971) , **Statistical Principles In Experimental Design** .NewYork :McGraw-Hill(2nd ed.)